Chapter Eight

LANGUAGE AND THINKING ABOUT THOUGHTS

At various points in this book we have examined different conceptions of the relation between linguistic and non-linguistic thought, and correspondingly of the scope and limits of non-linguistic thought. These different conceptions fall naturally into two broad groups. At one extreme we find those views that seek to drive as much of a wedge as possible between linguistic and non-linguistic thought. In Chapter Three, for example, we considered the minimalist conception of non-linguistic thought, according to which the thoughts of creatures that are not language-users cannot have either compositional structure or determinate contents. According to proponents of the minimalist conception, non-linguistic creatures are capable only of an etiolated and imagistic type of thinking, the vehicles of which are "spatial images superimposed on spatial perceptions". If the minimalist conception is correct then it will provide us with a very clear sense of the differences between linguistic thought and non-linguistic thought. So too would a view that goes naturally with the minimalist conception – the view, namely, that considerations of sense and mode of presentation can get no grip in the absence of language. But much of this book has been devoted to arguing that, initial appearances to the contrary, there is indeed room at the non-linguistic level for the ascription of thoughts that have determinate contents, compositional structure – and, more strongly, that such thoughts are indispensable for the purpose of

psychological explanation. So where does this leave us? Are there any genuine differences at all between linguistic thought and non-linguistic thought?

The principal claim of this chapter will be that there is an important class of thoughts, which is in principle unavailable to non-linguistic creatures. In the first section I will briefly discuss some of the claims that have been made about how language can function as a cognitive tool. It will emerge that many of these functions do not actually require a full-fledged language. If language is to have a distinctive, unique and essential cognitive function, so that the difference between linguistic and non-linguistic cognition is qualitative rather than merely quantitative, then we must look elsewhere. In 7.2 I develop my own account of the contribution that language can make to cognition – and explain why it is a contribution that can only be made by language. The central claim will be that all thinking that involves intentional ascent (roughly, all thinking that involves thinking about thoughts) requires the capacity for semantic ascent (roughly, the capacity to think about words). In the next chapter we will go on to explore the practical implications of this result.

8.1 Language as a cognitive tool

Andy Clark has usefully plotted out six fundamental respects in which language can function as a tool enhancing, extending and facilitating thought and cognition (Clark 1998).[1]

[1] As we will see below, these six respects do not exhaust Clark's conception of the contribution that language can make to cognition.

(1) Memory augmentation The acquisition of a public language offers us powerful means of systematically storing data (not simply in written language, but also in communicated oral traditions).

(2) Environmental simplification Applying linguistic labels is one way in which the perceived environment can be broken down into persisting objects and properties.

(3) Coordination Language permits the mutual control of attention and resource allocation in coordinated activities.

(4) Transcending path-dependent learning The learning patterns of linguistic creatures are not constrained by the particular cognitive paths they have followed since linguistic communication allows ordinary learning patterns to be circumvented and short-circuited.

(5) Control loops Language allows us to create control loops for our future behavior (by writing down plans that might be too complicated to keep in one's head, for example) as well as to register and respond to the instruction and regulation of others.

(6) Data manipulation and representation "Extended intellectual arguments and theses are almost always the product of brains acting in concert with multiple external resources. These resources enable us to pursue manipulations and juxtapositions of ideas and data which would quickly baffle the un-augmented brain" (Clark 1998 p.173).

These are all, no doubt, important ways in which language possession serves to extend and enhance thinking. But they do not, I think, mark types of thinking which are in principle available only to language-using creatures. As a first step in plotting out the scope of non-linguistic thought it will be

useful to explore how several of the functions that Clark identifies can be carried out by non-linguistic processes.

By "non-linguistic process" I mean one or other of two things. On the one hand a process can be straightforwardly non-linguistic in virtue of not relying upon any symbols at all.[2] On the other a process can be non-linguistic even though it is dependent upon symbols, simply because it does not rely upon the right sort of symbols. Of course, to describe a process as non-linguistic in this second sense is hostage to a principled distinction between linguistic and non-linguistic symbol systems and I shall have some more to say about that shortly. First, though, let us consider how far one can get with the prosthetic functions of language at an entirely non-symbolic level.

The mutual control of attention and resource allocation in coordinated social activities does not require the intervention of language. Coordination requires a degree of communication, but this communication can perfectly well be non-symbolic (and hence, uncontroversially, non-linguistic). It is well known, for example, that human infants engage from a very early age in sustained periods of coordinated activity with their caregivers. The process has aptly been called one of *affect attunement* (Stern 1985). It is a process of exploring and communicating emotional states through changes in facial expression, vocalizations and gesture.

> Affect attunement is the performance of behaviors that express the quality of feeling of a shared affect state without imitating the exact behavioral expression of the inner state. . . The

---

[2] I take it as uncontroversial that no system of communication can count as linguistic unless it involves symbols. The characteristic of symbols is that they are arbitrary as opposed to iconic signals (see further in the main text below and also n.3 in Chapter Two).

> reason attunement behaviors are so important is that true imitation does not permit the partners to refer to the internal state. It maintains the focus of attention upon the forms of the external behaviors. Attunement behaviors, on the other hand, recast the event and shift the focus of attention to what is behind the behavior, to the quality of feeling that is being shared. (Stern 1985, 142)

In affect attunement emotional states are communicated by being shared. The infant's participation in this sort of communication is clearly intentional (as, of course, is the care-giver's), and the process of affect attunement is equally clearly a type of coordinated activity in which there is a mutual control of resource allocation. Nonetheless, the process is not one of symbolic communication, because facial expressions and vocalizations are not *symbols* for the emotional states that are being communicated. The link between a scowl and a feeling of disagreeableness, for example, is expressive rather than symbolic. The behavioral manifestations of emotions and feelings cannot be divorced from the emotions and feelings that are being manifested. That is to say, the link between them is not arbitrary and conventional, which is the essence of the symbolic.

A further example of coordination without symbols comes with the well-studied dances carried out by worker honeybees (von Frisch 1967).[3] Although the details of the dances vary among the seven known species of the genus Apis, the basic principles are similar. Worker bees returning to the nest site after successful foraging land on the horizontal top of the nest and perform what has come to be known as a waggle dance to inform their fellow workers of the location of the food source. In the waggle dance the bee flies in a figure-of-eight pattern, moving its abdomen back and forth laterally on the straight line in between the two circles. Although there is of course a degree of error, the waggle dances communicate

[3] The bee dances have also received attention from philosophers. See Bennett 1964.

information about the direction, distance and desirability of the food source. Each of these three dimensions of variation is correlated with a dimension of variation in the dance. The angle of the dance relative to the position of the sun indicates the direction of the food source. The duration of a complete figure-of-eight circuit indicates the distance to the food source (or rather the flying time to the food source, because it increases when the bees would have to fly into a headwind). And the vigor of the dance indicates the desirability of the food to be found. Variation in all three environmental features is preserved in the relevant dimensions of the bee dance.

The bee dance is an even clearer example than affect attunement of coordination in resource allocation – and, as in the affect attunement case, it does not seem appropriate to describe the coordination as resting upon symbolic communication. The bee dance is non-symbolic because it does not involve arbitrary signals. To appreciate the point, and in particular the distinction between arbitrary and iconic signals it is worth making a small diversion into the theory of signals. There can be no transfer of information without rules for the encoding and decoding of signals (Green and Marler 1979). Viewed in broad and abstract terms, the rules for the encoding of a signal can be viewed as rules governing the transformations that lead from stimuli received at the sensory surfaces to signaling behavior. The simplest such rules will be when there exist deterministic links from input to output, as when the signaling behavior is what is sometimes called an 'instinctive behavior pattern' or an 'innate releasing mechanism'. Newly hatched herring gulls, for example, are particularly sensitive to the sensory input correlated with the length, movement and coloration of the adult herring gull's bill. A transformation rule maps this onto the chick's characteristic signaling behavior, vigorous pecking at the adult's bill tip (Tinbergen 1973). In this case the decoding signal is equally straightforward. The adult herring gull responds by feeding the chick. More complex signaling behaviors might involve a series of rules governing a series of transformations. In affect attunement, for example, there will be rules governing the extraction of various qualitative and quantitative aspects of vocal exclamations and gestures by the partner in the exchange; rules mapping

those aspects onto an amodal 'common currency'; and rules which then generate responsive gestures or vocalizations (Stern 1985).

In an iconic signal system there will be, for each dimension of variation in the relevant environmental feature (the distance of the food source in the honeybee case, for example), a single transformation rule mapping variation in that parameter onto variation in the relevant dimension of the signal. The receiver of the signal will decode the signal by an inverse mapping.[4] Signs are arbitrary, as opposed to iconic, when no such general and continuous rules exist to map variations in environmental parameters onto variations in signals. Each member of the signal set is linked to a particular environmental feature by specific assignments that can only be characterized one by one. Within the framework set by the distinction between arbitrary and iconic signals, it seems relatively uncontroversial, first, that the bee dances are iconic rather than arbitrary signaling practices and, second, that no practice that does not involve arbitrary signals can plausibly be described as linguistic.

The bee-dances, then, seem clearly to illustrate how coordination in resource allocation and communal activity can be achieved at the non-linguistic level. They also illustrate how control loops can be created

---

[4] The definition of iconic signals just posed is significantly different from that suggested by Charles Sanders Peirce, who is usually credited with having first formulated the distinction between iconic and arbitrary signals (or rather, in his terms, between iconic and symbolic signs). Peirce defines an icon as a "Sign that represents its Object in resembling it" (Peirce 1991, p.270). He does not impose any requirement that continuous variation in the object be matched by continuous variation in the signal, and consequently takes as iconic signs that would not count as iconic on the criteria just outlined. For example, he holds that linguistic predicates are iconic, as well as logical and mathematical proofs. Peirce's theory of signs is discussed in Chapters Four and Six of Hookway 1985.

without language and, moreover, how path-dependent learning can be transcended. The whole point of the bee dances is that the vast majority of worker bees do not need to seek the food for themselves. The worker bees that remain in the hive can profit from the explorations of a small number of bees to circumvent the standard search procedures.

Nor is this the only way in which the path-dependence of learning can be circumvented without recourse to language. Imitative learning is perhaps the most fundamental way of overcoming path-dependence through permitting the social transmission of knowledge. Imitation is extremely widespread in the animal kingdom and present in human infants more or less from birth (Meltzoff and Moore 1977, 1983). There are several different types of imitative learning, each of which offers a different way of understanding the possibility of path-independent learning. The most straightforward form of imitative learning is simple mimicry, of the sort that can be seen in the social transmission of bird songs and in neonatal imitation of facial expressions. In many species, individual birds deprived (e.g. by neo-natal deafening or by being raised in isolation) of the opportunity to imitate the songs of conspecifics develop either only very rudimentary version of the standard songs or radically abnormal songs (see, e.g., Marler 1970). At a more sophisticated level we can identify various types of mimetic learning, of different degrees of sophistication. It is known from studies of chimpanzee tool use, for example, that chimpanzees are skilled at learning about the dynamic properties of objects from observing them being manipulated by others (Boesch and Boesch 1992). There is some debate about whether tool-using chimpanzees are actually imitating patterns of behavior (Tomasello 1996), as opposed to picking up the dynamic affordances of objects and learning about features of the environment, but these clearly illustrate some form of the social transmission of knowledge.

The cognitive significance of mimetic learning within phylogeny has been recognized by Merlin Donald in his speculative reconstruction of the emergence of the modern mind (Donald 1991). Donald

identifies three principal stages in the development of hominid cognition. The distinctiveness of his account of hominid evolution comes with his suggestion that a long period of what he terms *mimetic culture* served as the bridge between the episodic form of cognition characteristic of the apes and earliest hominids and the language-based culture that emerged after the transition from the Middle Paleolithic to the Upper Paleolithic. Donald associates mimetic § culture with a wide range of social cognitive phenomena that collectively served as an adaptive foundation for the evolutionary emergence of language. The enormous development in techniques of tool manufacture characteristic of the Middle Paleolithic period must have been linked with new means for the transmission of information across a community. From what we know of contemporary hunter-gatherer societies the process of tool manufacture is very much a social process, and clearly dependent upon successful communication among the participants (Reynolds 1993). If we follow the majority of archeologists in dating the emergence of linguistic communication after the Middle Paleolithic (Bickerton 1990, 1996, Liebermann 1984) then it is hard to see any alternative to group mimetic learning as the crucial cognitive mechanism underpinning complex tool manufacture. Mimesis can also plausibly be viewed as the foundation for the social expression of emotion – again a form of non-linguistic communication that circumvents the path-dependence of language.

Returning to Clark's list of the instrumental cognitive functions served by language, we have already seen in some detail that at a very basic level the function of environmental simplification does not require any sort of communication, symbolic or otherwise. Human infants are born able to parse the perceived environment into object-like segments that are taken to obey certain basic higher-order physical principles. This is probably not enough, however, for the type of environmental simplification that Clark intends. His view, I suspect, is that the environmental simplification that language provides applies to a perceived environment that is already parsed into objects or object-like entities. Language does not enable us to perceive an environment composed of discrete and continuously existing objects in the way that

many philosophers have argued.[5] What it allows us to do is to impose a simplified pattern that will allow us to make sense of the discrete and continuously existing objects that we encounter in perception and action.

Even when the function of environmental simplification is understood in this sense, however, it still seems clear that it can be achieved at the non-linguistic level. It is at this point that it becomes pressing to say something about quite how "linguistic" and "non-linguistic" are being understood here. There has been considerable discussion among philosophers, linguists and psychologists as to what constitutes a language. Savage-Rumbaugh has argued, for example, that language emerges with the appearance of symbolic communication (Savage-Rumbaugh 1986). Most linguists, on the other hand, think that some form of grammar and syntactic structure is essential for a symbol system to count as a genuine language (Chomsky 1980). The majority of philosophers have followed the linguists in this respect (Bennett 1976),

---

[5] Consider, for example, the following two passages from Michael Dummett's book on Frege. "Our apprehension of reality as decomposable into discrete objects is the product of our application to an originally unarticulated reality of the conceptual apparatus embodied in our language" (Dummett 1973, 505). And: "Our ability to discriminate, within reality, objects of any particular kind results from our having learned to use expressions, names or general terms, with which are associated a criterion of identity which yields segments of reality of just that shape: we can, in principle, conceive of a language containing names and general terms with which significantly different criteria of identity were associated, and the speakers of such a language would view the world as falling apart into discrete objects in a different way from ourselves. . . For Frege, the world does not come to us articulated in any way; it is we who, by the use of our language (or by grasping the thoughts expressed in that language), impose a structure on it" (Dummett 1973, 503-504).

although David Lewis has proposed in a rather more inclusive spirit that languages are simply functions from expressions onto meanings (Lewis 1983). Confronted with this it is hard not to be pulled in two directions. On the one hand it is tempting to think that there is little point in trying to arbitrate in this debate. It is unclear what criteria we should use to judge any of these proposals. If we are trying to capture the essence of human language then it is clear where the answer lies - but our project is broader than that. The ordinary meaning of the word 'language' is too vague to help us. After all, ordinary language allows us to talk freely of the language of architecture and the language of the body. But, on the other hand, it seems impossible to investigate the nature and scope of non-linguistic thought without a demarcation line between the linguistic and the non-linguistic.

I will adopt middle way between these two extremes. A genuine language must allow for the formation of complex symbols from simple symbols. The essence of language is the combination of symbols with each other to express thoughts, taking thoughts to be complex entities that can be assessed for truth or falsity. The possibility of truth appears only with complex symbols. It emerges only when a state of affairs is asserted to be the case and this requires, as a bare minimum, the combination of a naming symbol and a predicate symbol. A lexicon of simple symbols will not support communicative utterances assessable for truth or falsity (although they may, of course, be adequate to perform speech acts such as issuing injunctions). So, complex symbolic systems allow the formation of complex symbols, symbols whose meaning is determined by the meaning of the symbols of which they are composed. That is to say, complex symbol systems display what is often known as *compositionality*, the possibility of recombining the members of a finite number of simple symbols to produce a range of complex symbols. This characteristic of complex symbol systems is frequently stressed in both philosophical and psychological discussions of language. It tends to be taken as a unitary phenomenon, but compositionality can be displayed in two fundamentally different ways, corresponding to two fundamentally different types of complex symbol systems.

In the first category, the category of sequentially complex symbol systems, complex symbols are formed in a purely additive or successive manner. Correlatively, understanding a sequentially complex symbol is a matter of successively understanding the simple symbols that make it up. A good example of such a sequentially complex symbol system comes with the communication systems spontaneously developed by very young children born deaf but whose parents do not use sign language to them. Such children tend without instruction to develop elementary signing systems employing combinations of two signs (Goldin-Meadow 1979). The two-sign combinations typically involve a pointing gesture combined with an action sign. An example might be pointing at a door combined with a turning gesture to signify that it should be opened. Pygmy chimpanzees, or bonobos (Pan paniscus) have acquired similar abilities (Greenfield and Savage-Rumbaugh 1990). It is clear that we are dealing here with the formation of complex symbols. In the given example it is the formation of a complex symbol communicating an injunction – the injunction to open the door. But the mode of formation is extremely simple. The syntactic rules of the symbol system are such that the action sign (the turning gesture) can only be conjoined with a sign that picks out an object. There is no way of using the syntactical rules to build up further complex symbols.

In what might be termed a hierarchically complex symbol systems, on the other hand, the logical and semantic relations between the components of a complex symbol cannot in every case be read off from the sequential ordering of simple symbols. The basic feature of a hierarchically complex symbolic system is that it possesses a hierarchically organized compositional semantics operating in such a way that understanding a complex symbol will not always be a matter of successively understanding the simple symbols that make it up. To get from a sequentially complex to a hierarchically complex symbol system we must lift the basic restriction that complex symbols can only be formed by combining a predicate symbol and the appropriate number of naming symbols. Hierarchically complex symbol systems possess a variety of mechanisms that allow the formation of complex symbols. Operators analogous to the definite

and indefinite articles in English will allow the formation of complex names (i.e. definite and indefinite descriptions), while reiterable operators that apply to complete sentences will allow the formation of complex symbols compounded from sentences. The logical constants are a case in point, but so too are the operators indicating possibility and necessity and indeed tense. I will have more to say about these in the next chapter (particularly in sections 8.4 – 8.7).

Returning to the matter in hand, it seems very plausible that the function of environmental simplification identified by Clark can be carried out by “labeling” environmental features with a simple symbol system. Environmental simplification does not require complex symbols (although of course it would be greatly enhanced by a complex symbol system), and hence does not require a language. A classic example of how this might work (and one that provides a further example of how the instrumental functions we have already considered can be effected at the non-linguistic level) comes with the alarm calls of vervet monkeys (Cheney and Seyfarth 1990). Vervet monkeys have three alarm calls, each geared to a different category of predator – eagle, leopard and snake. When vervets hear a particular alarm call they do not display an indiscriminate fear or avoidance response, but rather behave in ways appropriate to the predator to whose presence they have been alerted. When they hear the eagle alarm call, for example, they look up and scan the sky. Vervets also have two further calls, the "wrr" and "chutter" calls, which communicate the nearby presence of strange groups of vervets. According to Cheney and Seyfarth, vervet monkeys are more likely to issue alarm calls when there are close kin in the vicinity (1990 Ch. 5). Something similar holds of the food signals issued by macaque monkeys, who seem to scan for members of their social group before issuing a food signal (Hauser and Marler 1993). Even if we assume that the vervet alarm calls are symbolic, they do not constitute a complex symbol system, since there is no scope for the formation of complex symbols. Yet it is clear that the alarm calls are serving a function of environmental simplification, and indeed permit the creation of control loops for regulating each individual’s future behavior and the behavior of other members of the group.

Nor is it only in primates that we find such behaviors. A classic ethological example of environmental simplification and control loop creation comes with the ways in which different species of birds hide caches of seeds at specific locations. Clark's nutcrackers are known to deploy multiple cues to identify the locations of nut caches. Experiments have show that they are able to reidentify nut stores even when experimenters manipulate the most obvious landmarks (Balda and Turek 1984). In creating food stores birds both simplify the environment in terms of markers which allow them to reidentify the caches, and create control loops for future behavior in terms of a structure of inter-cache trajectories which will determine their movements when retrieving food.

As far as Clark's original list of six instrumental functions is concerned, therefore, it looks as if we can make sense of at least four of them operating at the non-linguistic level. And the two that are left (memory augmentation and data manipulation) do not seem very likely candidates for marking a distinctive type of cognition only made available by the acquisition of language. There are all sorts of ways in which non-linguistic creatures can augment their limited memory capacities, some of which we have already considered in thinking about how path-dependent learning might be transcended at the non-linguistic level. And data manipulation is a characteristic of all thought. The only differences between linguistic and non-linguistic thought in either of these dimensions are likely to be differences of degree. It would seem, therefore, that if there indeed are any fundamental differences of type between linguistic and non-linguistic cognition they will lie elsewhere. In the next section I will turn to a fundamentally different type of argument for the distinctiveness of linguistic cognition.

8.2 Intentional ascent and semantic ascent

In the same article in which he puts forward the six instrumental functions of language that we considered in section 7.1 Andy Clark offers the following conjecture about the role of language in underpinning the distinctiveness of human cognition.

> Perhaps it is public language that is responsible for a complex of rather distinctive features of human thought – viz., our ability to display *second-order cognitive dynamics*. By second-order cognitive dynamics I mean a cluster of powerful capacities involving self-evaluation, self-criticism and finely honed remedial responses. Examples would include: recognizing a flaw in our own plan or argument, and dedicating further cognitive efforts to fixing it; reflecting on the unreliability of our own initial judgements in certain types of situation and proceeding with special caution as a result; coming to see why we reached a particular conclusion by appreciating the logical transitions in our own thought; thinking about the conditions under which we think best and trying to bring them about. The list could be continued, but the pattern should be clear. In all these cases we are effectively thinking about our own cognitive profiles or about specific thoughts. (Clark 1996, 177)

He explains how language makes these types of thought available in the following terms.

> It is easy to see in broad outline how this might come about. For as soon as we formulate a thought in words (or on paper), it becomes an object both for ourselves and for others. As an object it is the kind of thing we can have thoughts about. In creating the object we need have no thoughts about thoughts – but once it is there, the opportunity immediately exists to attend to it as an object in its own right. The process of linguistic formulation thus creates the stable structure to which subsequent thinkings attach. (Clark 1996, 177)

It seems to me that the basic idea Clark is putting forward is essentially correct. My aim in this section of the chapter will be to formulate the basic idea more precisely and to provide an argument to back it up.

Clark's explanation of the role language plays in second-order cognitive dynamics is incomplete, at least for our purposes. For one thing, all he really offers is an account of how, given that we have language, we are able to engage in second-order cognitive dynamics – whereas what we need is an argument that second-order cognitive dynamics can only be undertaken by language-using creatures. This links up directly with a more fundamental worry. The natural way to derive an argument for the necessity of language from Clark's suggestions would be to claim that language is required for thinking about our own thoughts. But this claim is hardly uncontroversial. It would be instantly denied, for example, by language of thought theorists. It is an integral part of the language of thought hypothesis that reflexive thinking is available in the language of thought. Once again we find ourselves caught in the familiar grey area between public languages and the language of thought. We need to investigate, not simply whether some form of language is required, but what type of language that must be.

The basic premise from which we need to start here is that reflexive thoughts can only be possible if the target thoughts have vehicles that allow them to be the objects of further thoughts. The question, then, is what form these vehicles must take. Broadly speaking the candidates fall into two broad categories. They might be either personal-level vehicles or subpersonal vehicles. Any argument to show that reflexive thinking requires the target thoughts to be vehicled in a public language will have to establish, first, that the vehicles must be at the personal level and, second, that the only available vehicles at the personal level are public language sentences.

The plausibility of the first step in the argument emerges when we reflect that reflexive thinking will paradigmatically involve a direct and conscious cognitive access to the target thoughts. As the first

passage quoted from Clark at the beginning of this section makes clear, typical examples of reflexive thinking might include evaluating evidential and inferential relations between thoughts. Such reflexive thinking involves having the target thoughts in mind – entertaining them consciously and considering how they relate to each other logically and evidentially. Yet, it seems clear that we do not stand in the appropriate sort of direct and conscious cognitive access to subpersonal states. Second-order cognitive dynamics is a matter of the conscious regulation and policing of one's own thoughts - and we do not consciously regulate and police sentences in the language of thought. Of course, it might be the case that certain types of hypothesis testing and refinement do take place at the subpersonal level. Something like this happens, according to Fodor, when we learn a language. Nothing I say is incompatible with that proposal, since my claim is simply that such processes would not count as instances of second-order cognitive dynamics. Nonetheless, one might want to question the proposal on other grounds, such as the availability of alternative models of language acquisition, particularly those developed within the connectionist tradition (for a survey of recent work see McLeod et al. 1998).

It is hard to see how a defender of the language of thought hypothesis could maintain that we do have conscious access to sentences in the language of thought, given that the language of thought hypothesis is a hypothesis about subpersonal cognitive architecture, not about the medium of conscious thought. It is generally accepted that subpersonal states are inferentially insulated from the conscious processes of cognitive evaluation and self-criticism. This is precisely the distinction between the personal and subpersonal levels (see, for example, Stich 1969). *A fortiori*, therefore, it follows that the vehicles of the

thoughts which are the objects of what Clark calls second-order cognitive dynamics cannot be sentences in a subpersonal language of thought.[6]

There are versions of the language of thought hypothesis, however, that are not put forward purely as hypotheses about subpersonal cognitive architecture and hence that are not directly blocked by this line of argument. Some authors have suggested that there might be a modular internal language in which thoughts are encoded for conscious consideration. This general approach would sit well with (although neither entails nor is entailed by) higher-order thought theories of consciousness (Rosenthal 1991). It seems to me, however, that this proposal lacks phenomenological plausibility. All the *propositional* thoughts which we consciously introspect, whether thoughts which come unbid into our minds or thoughts which we consciously instigate and consider in solving a problem, take the form of sentences in a public language. We do, of course, engage in various types of non-sentential thinking – what, in Chapter Three, I termed thinking-how in contrast to thinking that – and it is only mistaking non-propositional thinking-how for propositional thinking that which leads to the view that we can be aware of "pure"

---

[6] Nothing I say is intended to be incompatible with the substantive claims made by the theorists in the language of thought tradition. In particular, it may well be the case (as Fodor suggests) that the understanding of public language sentences involves in some sense translating them into the language of thought. My point is simply that reflexive thinking (in the sense in which I am understanding it, namely, as involving, e.g., the evaluation of epistemic links between propositions) is directed at the public language sentences rather than at the sentences in the language of thought which give their meaning.

thoughts without any sentences featuring among the contents of introspection.[7] There are certain types of problem that we solve by manipulating mental images and exercising the visual imagination. And we are, of course, conscious of bodily sensations, emotional feelings and other such qualitative states (although these are not properly described as types of thinking at all). But we are not, I think, ever conscious of *propositional* thoughts that do not have linguistic vehicles. When we are conscious of propositional thoughts we are conscious of imaged sentences.[8] What we introspect when we introspect our propositional thoughts in the manner required for the processes of second-order cognitive dynamics is inner speech.

---

[7] The point was well put by Wittgenstein: "What happens when we make an effort – say in writing a letter – to find the right expression for our thoughts? – This phrase compares the process to one of translating or describing: the thoughts are already there (perhaps were there in advance) and we merely look for their expression. This picture is more or less appropriate in different cases. – But can't all sorts of things happen here? – I surrender to a mood and the expression *comes*. Or a picture occurs to me and I try to describe it. Or an English expression occurs to me and I try to hit on the corresponding German one. Or I make a gesture and ask myself: What words correspond to this gesture? And so on." (Wittgenstein 1953 §335). For further discussion of Wittgenstein's complex views on the relation between language and thought see Budd 1989 Chs. 5 and 6.

[8] Peter Carruthers, who proposes that all domain-general cognition consists in the formation and manipulation of linguistic representations at the level of what Chomsky terms logical form, nonetheless thinks that we cannot be conscious of these "stripped down" linguistic representations (Carruthers 1996, forthcoming). We are only conscious of sentences with the full complement of natural language phonological and structural features from which the level of logical form is an abstraction.

This brings us to the second step of the argument. What needs to be shown is that public language sentences are the only possible personal-level vehicles for thoughts that are to be the objects of reflexive thinking. The conclusion so far is that sentences in the language of thought are not appropriate vehicles. The only way to proceed is by elimination – showing that there are no viable alternatives to public language sentences. How else might the representation relation work? There seem to be two possibilities. On the one hand representation might be secured symbolically through the complex symbols of a natural language (complex symbols being required since what are being represented are thoughts about states of affairs). A thought would be represented, therefore, through its linguistic expression and would appear as a potential object of thought *qua* linguistic entity. On the other hand representation might be secured in an analogue manner, through some kind of pictorial model. On this conception of the vehicles of thought, which we find developed in different ways in mental models theory in the psychology of reasoning (originally proposed in Craik 1943 but most comprehensively developed in Johnson-Laird 1983) and in the conception of mental maps put forward by Braddon-Mitchell and Jackson (Braddon-Mitchell and Jackson 199?), the vehicle of a thought is a pictorial representation of the state of affairs being thought about.[9]

---

[9] I will count mental models theory as a conception of pictorial models even though the two notions cannot be straightforwardly mapped onto each other. It is true that mental models, as proposed by Johnson-Laird, are intended to be semi-perceptual states that resemble the situations they represent. Nonetheless, there is a crucial ambiguity in mental models theory. On the one hand the manipulation of mental models is supposed to take place in working memory, the contents of which are generally thought to be open to conscious access and report. On the other hand, it seems clear that introspection will have little role to play in deciding the issue between mental logic and mental models as far as the psychology of reasoning is concerned. To the extent that mental models theory is a theory of the

The idea of structural isomorphism is at the heart of both theories. Both mental models and maps are spatially isomorphic with what they represent. The relations (or at least some of them) holding between elements of the mental model/map can be mapped on to the relations holding between objects in the represented state of affairs. In this way representation is secured through the relations of exemplification and resemblance. The mental model/map represents a state of affairs by exemplifying that state of affairs' structure – that is to say, by itself possessing a structure which resembles (at some suitable level of abstraction) the structure of the represented state of affairs. It does so, however, in a way that does not have an independently identifiable structure corresponding to the state of affairs exemplified. Braddon-Mitchell and Jackson put the point clearly.

> There is no natural way of dividing a map at its truth-assessable representational joints. Each part of a map contributes to the representational content of the whole map, in the sense that had that part of the map been different, the representational content of the whole would have been different. Change the bit of the map of the United States between New York and Boston, and you change systematically what the map says. This is part of what makes it true that the map is structured. However, there is no preferred way of dividing the map into basic representational units. There are many jigsaw puzzles you might make out of the map, but no single one would have a claim to have pieces that were all and only the most basic units. (Braddon-Mitchell and Jackson p.171)

We need, therefore, to distinguish weak and strong senses in which a representational vehicle might be structured. In the weak sense there is structure whenever a structural isomorphism can be identified

---

subpersonal mechanisms of thought it will be correspondingly of less use as a theoretical account of how thoughts might be vehicled in a way that would allow them to be the objects of further reflexive thinking.

between the vehicle and what it represents. In the strong sense, however, structure requires the existence of basic representational units combined according to independently identifiable combinatorial rules. Natural language sentences (or for that matter sentences in the language of thought) are clearly structured in the strong sense, whereas mental maps/models only possess structure in the weak sense.

In mental models and maps the representation relation will be secured through some combination of isomorphic resemblance and exemplification holding primarily between the model/map as a whole and the represented state of affairs as a whole (and only derivatively at the level of the distinguishable elements of the map/model and state of affairs respectively). There are important questions to ask about the nature, and even the possibility, of resemblance and exemplification-based representation – questions that go back to early criticisms of the British Empiricists (and which have arisen more recently in criticisms of the picture theory of the proposition offered by Ludwig Wittgenstein in the *Tractatus Logico-Philosophicus*). I shall prescind from these, however. There is a fundamental problem, however, with either version of the analogue proposal as an account of how it is possible for thoughts to have vehicles that allow them to become the objects of further thoughts.

The problem is, in essence, that second-order cognitive dynamics involves sensitivity to the inferential relations between thoughts and we not yet have an understanding of how images can be inferentially connected to each other. The problem once again derives from the intimate relation between inference and structure explored in earlier chapters. There is a sense in which mental models and maps are structured, since they contain elements that can feature in further mental models/maps. Nonetheless, they do not seem to be structured in the right sort of way to permit the reflexive type of second-order cognitive dynamics under discussion. It will be helpful to take maps and mental models separately.

Let us start by considering how the mental maps approach might deal with second-order cognitive dynamics. A canonical example of second-order cognitive dynamics might be coming to see the evidential basis for a particular belief and then evaluating the inferential transition made on that basis. (This would be an example of what in section 9.1 I will term *reflective doxastic modification*.) An example might be the realization that one has made an overly rash inductive generalization, or a faulty deductive inference. It is perfectly easy to see how there could be some very basic forms of inferential transition between maps. Such transitions might be modeled on broadly associationist lines, and it is the possibility of such transitions that enables maps to serve as guides to action. What is not possible, however, is for such transitions to be understood and evaluated in terms of either deductive validity or probabilistic support. Those very features of maps (their analogue nature and structural isomorphism with what they represent) that make them so useful for guiding action serve to make them inappropriate for the type of inferential evaluation characteristic of second-order cognitive dynamics. In order for such evaluation to take place the maps must be interpreted in broadly propositional terms. We must interpret one map as expressing one proposition and the second as representing a further proposition, and then evaluate the inferential relations (be they deductive, inductive or probabilistic) between those two propositions. Once again, our only understanding of how to do this rests upon the two propositions being linguistically formulated.

Braddon-Mitchell and Jackson do not directly address this issue, but they do offer the following explanation of how maps can evolve over time in what is clearly intended to be an analogy with inferential transitions between linguistically vehicled representations.

> Maps are physical entities whose structure can govern the way they evolve over time. When cartographers update maps or put two maps together to make one that incorporates all the information in a single map, these operations are governed in part by the structures of the

> maps they are working on. And in order to find a target, rockets use a kind of internal map that gets continually updated as new information comes in. In these rockets, later maps are causal products of earlier maps plus what comes in via the rocket's sensors. Hence map theorists can tell an essentially similar story to language of thought theorists about how thoughts evolve over time as a function of their propositional objects. (Braddon-Mitchell and Jackson p.173)

There is a fundamental disanalogy here, however. The issue is not really about how thoughts evolve over time. In a very important sense individual thoughts quite simply do not evolve over time. It is systems of thought which evolve, and they do so as a function of the inferential relations between the thoughts that compose them. These inferential relations hold between distinct thoughts and nothing that Braddon-Mitchell and Jackson say in this short passage give us any way of understanding how we should understand inferential relations between distinct thoughts at the level of mental maps. The process of combining maps has only very limited analogies with the process of inferring one thought from another. We do not have, for example, any idea what a conditional map might look like – and consequently little understanding of how conditional reasoning might take place at the level of mental maps. *A fortiori*, therefore, there seems no sense in which we can understand second-order cognitive dynamics as applying to the inferential transitions between mental maps.

The situation is somewhat more complicated with mental models theory, given that mental models have been expressly put forward as a descriptive account of how reasoning takes place. How can mental models not be suitable targets for the type of reflexive thinking characteristic of second-order cognitive dynamics, given that mental models are explicitly proposed as providing a unified account of deductive, probabilistic and modal reasoning (Johnson-Laird 1999). It is important to be clear, however, about the precise claims of mental models theory. Mental models theory is proposed as an alternative to the so-

called mental logic theory (Rips 1994), according to which reasoning is an exclusively syntactic matter grounded in formal rules of inference.[10] The key idea of mental models theory is that arguments are evaluated by the construction of mental models of the relevant premises. An argument is judged to be deductively valid if the conclusion holds in all the constructed models of the premises, probabilistically valid if it holds in most of the models, and so forth. The claim made is that this model of reasoning provides a better explanation of the patterns discovered in the experimental study of how subjects reason than the idea that reasoning involves the manipulation of sentential representations according to formal rules. Some of these patterns involve systematic susceptibility to formal fallacies, which it is obviously difficult to explain on the mental logic approach. Others are simply patterns in the time taken to carry out certain inferences, which advocates of the mental models approach claim is directly correlated with the number of models which the reasoner needs to construct (Johnson-Laird and Byrne 1991).

Mental models are indeed supposed to be structurally isomorphic to the states of affairs described in the premises. The important point, however, is that, like mental maps, their structure is derivative. It is derived from the premises that they are modeling. The models are constructed from constituents and properties that feature in the premise being modeled. And those premises are of course linguistic entities. It would be a mistake to think that mental models theory construes inference in terms of transitions between mental models – any more than a model-theoretic approach to the sentence calculus construes inference in terms of transitions between truth tables. Just like any other theory of reasoning, mental models theory construes inference as a matter of transitions between sententially encoded propositions. What is distinctive about it is that it construes those transitions between sententially encoded propositions

---

[10] The difference between the mental logic and mental models theories is frequently compared to that between proof-theoretic and model-theoretic approaches in logic.

as taking place in virtue of relations between analogue representations of the states of affairs portrayed in those sententially encoded propositions (as opposed to formal relations holding between the syntactic structures of the relevant sentences). The point, in essence, is that mental models theory is not a genuine alternative to the sentential conception of reasoning, since the whole idea of a mental model only makes sense within the framework of the sentential conception. Mental models theory offers a particular way of developing the sentential conception, not of supplanting it. Mental models are not the vehicles of inference, but rather, as their name suggests, models of those inferences.

By a process of elimination, therefore, we have reached the conclusion that thoughts can only be the objects of the type of reflexive thinking in which thoughts are the objects of thought if they have natural language vehicles. This is not, of course, to say that we cannot deploy mental maps and mental models. It is clear that we do, and it is highly likely that non-linguistic creatures do as well. Nor is it to say that mental models and mental maps cannot be the objects of thought. It seems clear that we can think about mental models and mental maps as well as think by means of them. The point is that we cannot use mental maps or mental models for thinking about thoughts in the manner demanded by second-order cognitive dynamics. Natural language sentences are the only proxies that will permit thoughts to function as the objects of thought in this manner. To put the matter in the form of a slogan, there can be no intentional ascent without semantic ascent. We think about thoughts through thinking about the sentences through which those thoughts might be expressed. The significance of this thesis depends, however, on the types of thinking that constitutively involve intentional, and hence semantic, ascent. We began this section with Clark's programmatic suggestion that only language-users can be capable of the types of cognitive self-criticism and self-monitoring that he terms second-order cognitive dynamics. In the next chapter I will work towards a more wide-ranging classification of the scope and limits of non-linguistic thought.