



PERGAMON

Language & Communication 22 (2002) 353–373

LANGUAGE  
&  
COMMUNICATION

www.elsevier.com/locate/langcom

# Thinking about thinking: language, thought and introspection

Peter Slezak\*

*Program in Cognitive Sciences, University of New South Wales, Sydney, NSW 2052, Australia*

---

## Abstract

I do not think that the world or the sciences would ever have suggested to me any philosophical problems. What has suggested philosophical problems to me is things which other philosophers have said about the world or the sciences.

(G.E. Moore, 1942, p. 14)

Peter Carruthers has made a vigorous attempt to defend the admittedly unfashionable doctrine that we think ‘in’ language, despite its displacement by something like Fodor’s ‘language of thought’. The idea that we think in language has considerable intuitive persuasiveness, but I suggest that this is not the force of good argument and evidence, but a familiar kind of introspective illusion. In this regard, the question of language and thought derives a more general interest, since the illusion is independently familiar from other notorious disputes in cognitive science such as the ‘imagery debate’. © 2002 Published by Elsevier Science Ltd.

*Keywords:* Language; Thought; Mentalese; Introspection; Imagery; Homunculus

---

## 1. Introduction: methodological preliminaries

In the wake of the ‘cognitive revolution’ the idea that we think ‘in’ language has been displaced by something like Fodor’s (1975) ‘language of thought’ (LOT) or ‘mentalese’ which is emphatically not seen by its advocates as a natural language.<sup>1</sup> The properties of natural language itself, such as syntactic structure and productivity, are taken to reflect, or to be derived from, those of mentalese. Hotly debated alternatives to a ‘language of thought’, such as connectionist or dynamical systems, even reject those features of internal representations such as structure and systematicity which are reflected in natural language. Accordingly, such views are even more

---

\* Tel.: +61-2-9385-1000.

E-mail address: p.slezak@unsw.edu.au (P. Slezak).

<sup>1</sup> But see Harman (1973) and Devitt and Sterelny (1987) for closely related views.

remote from any conception of language as the medium of thought. Whatever their differences, these prevailing doctrines are all alternatives to a cognitive conception of language.

In view of its official unpopularity, it is surprising that the theory is still frequently discussed by its critics as if it were a live option deserving some gesture of acknowledgement, if only in order to refute it. Typical is Jackendoff's (1997) recent discussion which goes so far as to raise the issue irrelevantly in the context of a work on technical linguistics—openly admitting the oddity of this attention. Nevertheless, Jackendoff gives the issue serious treatment with a detailed response to the most common grounds for believing that natural language might be the medium of thought. I will suggest that the persistence of the cognitive view of language, and the evident need to keep refuting it, may be explained by the failure of both sides to address the crucial source of its illicit attractiveness. For example, sufficient evidence that Jackendoff, like Pinker (1994), misses the point is the fact that his rebuttals are not new, and have notoriously failed to lay the doctrine to rest.

In efforts to revive the 'cognitive conception' of language, Carruthers and Boucher (1998) acknowledge the dominance of the 'communicative conception'—namely, that language serves for communicating thoughts rather than as a vehicle embodying them. Accordingly, as a recent, strenuous attempt to defend the admittedly unfashionable cognitive thesis, the arguments of Carruthers (1996, 1998) deserve special attention (see Fodor, 1998). Carruthers concedes the kinds of arguments and evidence cited by Jackendoff (1997) and Pinker (1994), but tries to defend a carefully delimited thesis to which he thinks these do not apply. For his part, Davidson (1997) recently declares the case of Pinker and others as "flawed and the conclusions confused" saying further "The arguments for the existence of a language of thought prior to, or independent of, a socially engineered language are feeble" (1997, p. 20). However, I will suggest that these various defences are of interest mainly for the explicitness with which they reveal the notorious inadequacies of the doctrine that we think in language.

It is clear that, despite its systematic scientific shortcomings, the idea that we think in language has some persuasiveness which needs to be addressed. I will suggest that this persuasiveness is not the force of good argument and evidence in the usual sense, but a familiar kind of intuitive compellingness which is a seductive illusion. In this regard, the question of language and thought derives interest beyond the specific issues raised, since the seductive illusion afflicts inquiry into the mind quite generally and is familiar from other notorious disputes in cognitive science. The issue of language and thought is, therefore, a valuable case study of a pervasive mistake in reasoning about the mind or thinking about thinking. Indeed, there are grounds for suspecting that the doctrine of mentalese itself may not be free from the same mistake (Slezak, 2000, in press).

## 2. *Argumentum ad hominem*

In general, beyond a certain point, if a theory is thought to be implausible enough, we may seek to explain the *causes* rather than the reasons for holding it. For example,

we might cite distractedness, emotional state, intoxication, brainwashing or vested interests to explain why someone holds a sufficiently absurd or unlikely view. It is a version of this causal, ad hominem strategy which I adopt here. Instead of the evident futility of rehearsing the usual case against thinking in language, I simply assume it in order to ask the interesting question: If the case against thinking in language is correct, why should this case persistently appear so implausible? Paradoxically, this is to acknowledge a persuasiveness of the thinking-in-language view while also claiming it to be unworthy of being taken seriously.

This seeming paradox is reflected in the state of the controversy which is polarized between the extremes of those like Davidson (1975) and McDowell (1994) who take the thesis to be overwhelmingly obvious or even necessarily, conceptually true, and those like Magee (1997) who express incredulity that anyone might believe it except as a consequence of some pathology in their mental life. Resolving this paradox will be a central preoccupation of the present discussion.

With Magee (1997), I regard the thesis that we think in language as not merely false but extravagant, especially with its familiar corollary denying thought to animals. I take the evidence of thought without language (Weiskrantz, 1997; Wilkes, 1997) to be so overwhelming as to be beyond serious doubt. For the sake of argument, then, I take the falsity of the thesis that we think in language for granted in order to explore the cogency of reasons which have been offered in favour of it. My strategy has an important affinity with the question asked by David Armstrong (1980) about the intuitive implausibility of materialism as theory of mind. He suggests:

One thing that would greatly strengthen the Materialist case here would be the production of an *independently plausible* explanation of why Materialism is introspectively implausible. (1980, p. 50).

By analogy, I think that the case for mentalese and the communicative conception of language would be greatly strengthened by the production of an independently plausible explanation of why the view is introspectively implausible. The great virtue of Carruthers' (1996) defence of the cognitive conception of language is that he makes its crucial dependence on introspective intuition entirely explicit.

### 3. Pathology of inner life?

On the other side, too, it is by appealing to introspection that Magee (1997) expresses incredulity that anyone might entertain the doctrine of thinking in language. Magee remarks:

I understand the view: what I do not understand, and have never been able to understand, is how anyone can hold it. For I find (and I do not believe that I am differently constructed from others in this respect) that it is directly contradicted by my immediate experience. (1997, p. 78)

Magee suggests that the idea that all experience is capable of being expressed in words reveals a “staggering blindness” and “would be too preposterous to merit a moment’s entertainment” were it not for the fact that it is so widely held. He asks:

...how is it possible for so many undoubtedly clever people to assert, and believe in all sincerity, [that thought is in language]? ...How can they possibly believe what they say? What must their inner lives be like? ...Taking them to be sincere, which I do, they must be unlike most of the rest of us in the nature of their inner life and the way they experience it. (1997, p. 86)

Apart from its invalidity, as Magee himself recognizes (1997, p. 90), his particular *ad hominem* argument from introspection is a two-edged sword and may be turned against him. As he says, it is open to a linguistic philosopher to protest that it is Magee’s *own* inner life which is peculiar. The self-certifying nature of introspection means that there appears to be no way to adjudicate between these competing claims. However, the situation is not symmetrical. Even if Magee is exceptional, his case serves as a counter-example to any universal generalization that we think in language. Arguments about who is more typical are beside the point since it might be that people are simply different in their inner lives.

At best, however, Magee wins a Pyrrhic victory since his non-verbal introspections are no more explained by this fact than if they were in language. The bare avowals of inner life leave the fundamental explanatory question entirely untouched. Furthermore, Magee’s argument provides no answer to Carruthers (1996) who does not deny that we may think in other ways, but only insists that a centrally important part of human conscious thought is in language. Magee does not address the more fundamental question of the very *intelligibility* of the idea that we think “in” language (or anything else),—a question which had been raised by Ryle (1968) and which I wish to pursue presently. This question concerns the denial that we think in language in a different and stronger sense than the one Magee considers.

It is noteworthy that Magee’s incredulity and his assumptions about the defective inner life of rival theorists is not an isolated case. Leading researchers in the debate over visual imagery (Pearson et al., 2000) have circulated a questionnaire among other researchers inquiring about subjective experiences of imagery with a view to discovering possible correlations with theoretical stance. Like Magee, these researchers evidently believe that some deficit of mental life prevents others from appreciating the virtues of the favoured theory. However, vivid experience of visual imagery is consistent with *both* competing theories—Kosslyn’s (1994) pictorial account and Pylyshyn’s (1973) ‘tacit knowledge’ alternative. Seeing this mistaken assimilation of evidence with a particular theory helps us to understand why pictorialists have charged their rivals with denying the very *phenomena* of imagery to be explained. As we will see, in view of his dependence on introspection, it is not surprising that Carruthers makes the same charge against rivals of his thinking-in-language doctrine. Undeniably, introspective reflection can suggest that we are looking at pictures when we imagine, just as it suggests that we talk to ourselves when we think, but introspection may not reflect the underlying mechanisms giving rise to

them. Of course, in general it may be self-serving to claim that a rival theory is held only because of some intellectual defect in its adherents which blinds them to its obvious falsity. Nevertheless, in the present case, suitably refined, something like Magee's style of ad hominem argument may be deployed in a more subtle Duhemian (1906) form which is based on a theory of the instrument—in this case, introspection.

#### 4. Conceptual or constitutive conception

What exactly is meant by the claim that we think in a natural language such as English? The thesis permits more than one interpretation. In one version, the claim is that thought is “constitutively” or necessarily linguistic as a conceptual matter. Davidson (1975) argues for such a conceptual version of the claim. Among other philosophers holding this view, Carruthers and Boucher (1998, p. 5) identify McDowell (1994), too, who somewhat obscurely speaks of language as “a prior embodiment of mindedness” into which a human being is initiated (1994, p. 125), thereby “acquiring a mind” (1994, p. 126). For his part, Davidson claims to “show how thought depends on speech” and that “a creature cannot have thoughts unless it is an interpreter of the speech of another” (1975, p. 8,9).

Davidson's general *a priori* approach is one which still has some adherents (Woodfield, 1982 p. ix), but seems difficult to justify nowadays in the climate of cognitive science and a ‘naturalized’ approach to epistemology. Of his approach Davidson says:

...talk apparently of thoughts and sayings does belong to a familiar mode of explanation of human behaviour and must be considered an organized department of common sense that may as well be called a theory. One way of examining the relation between thought and language is by inspecting the theory implicit in this sort of explanation. (1975, p. 10)

That is, in order to understand the relation of language and thought we should consider what ordinary folk-psychology may say about it. Whatever the merits and status of folk psychology, it is difficult to see how it is any more likely to illuminate mental phenomena themselves than, say, the Aristotelian intuitions of folk-physics are likely to illuminate physical phenomena. One need not harbour eliminativist sentiments to doubt that conceptual analysis of folk psychology is the most useful source of explanatory insight about the mind.<sup>2</sup>

<sup>2</sup> We see exactly Davidson's approach echoed more recently by Woodfield (1982, p. ix) who says that “many of the most fascinating puzzles in the philosophy of mind have to do directly with everyday psychological idioms and the curious ways in which they work” and that the whole subject of philosophy of mind “is built upon a realization that philosophers can contribute more by investigating discourse about mental states than by investigating the mental states themselves”. Stich (1992) has carefully distinguished these fundamentally different enterprises and warned about confusing them. Nevertheless, McGinn (1997) avers “Philosophy, for me, is still anterior to science, and largely independent of it” (1997, viii). McDowell

Of course, in the nature of the case, the conceptual, constitutive versions of the doctrine that we think in language could not be refuted by the kind of empirical evidence cited by Jackendoff (1997), Pinker (1994), Weiskrantz (1997) or Wilkes (1997). However, I will suggest that, even for purportedly empirical versions of the thesis, this evidence is beside the point. Despite appearances and intentions, I will argue that the thesis is not a straight-forwardly empirical one (if there is ever any such thing), but rather a conceptual confusion. In its most common and persuasive formulation, the claim that we think in language cannot be coherently articulated. In this sense, ironically, the thesis cannot be established by conceptual analysis but may be dispelled by one.

For our present purposes, then, since Carruthers and Boucher (1998) seek to defend an empirical rival to Fodor's mentalese, we may happily endorse their comments rejecting any a priori conception and recommending that "all claims to the conceptual, or logical, involvement of language in thought can be dropped". Instead, we may consider the merits of any claim about language and thought as an empirical matter which they refer to as "natural" necessity (1998, p. 12).

## 5. Thoughts and their ascription

Before proceeding in this direction, however, it is worth noting that Davidson does, in fact, give some reason to think that his position may be more subtle than the kind of ordinary language a priorism just indicated. He seems to be concerned not only with the question of language and thought itself, but with the conditions and justification for the *attribution* of thought—an entirely different matter as Devitt (1984) has clearly pointed out in a different context. Thus Davidson says "...the attribution of desires and beliefs (and other thoughts) must go hand in hand with the interpretation of speech...But it remains to say...why the attribution of thought depends on the interpretation of speech." (1975, p. 15). Davidson remarks:

These considerations will probably be less persuasive to dog lovers than to others, ... At best what we have shown, or claimed, is that unless there is behaviour that can be interpreted as speech, the evidence will not be adequate to justify the fine distinctions we are used to making in the attribution of thoughts. (1975, p. 16)

Davidson seems to be concerned with the problem of how to characterize beliefs in dumb creatures using the distinctions available to us, since "Our manner of attributing attitudes ensures that all the expressive power of language can be used to make such distinctions" and "the intensionality we make so much of in the attribution

---

(1994, p. 89), too, expresses open disdain for scientifically informed theorizing which he labels "bald naturalism" and "unreflective scientism" which he sees as an immunity to certain problems which "ought not to be mistaken for an intellectual achievement". McDowell himself is not likely to be guilty of any "unreflective scientism" since the closest he gets to any relevant science is a passing footnote mention of the Müller-Lyer illusion (1994, p. 11).

of thoughts is very hard to make much of when speech is not present” (1975, p. 15,16).<sup>3</sup> Glock (1997, p. 167), too, seems to think that we cannot individuate or describe thoughts unless they are *capable* of being expressed in language. However, this is a conclusion which would preclude any scientific understanding of the psychology of other creatures. The concern is presumably a matter of theoretical vocabulary and what may be the best technical language for characterising the phenomena of interest, but it is difficult to see why this question is in any way affected by the difficulty of articulating the thoughts of animals in the vocabulary of our own natural language.

## 6. Understanding: intelligibility or explainability?

The alleged problem we have just noted seems to arise from a tacit assumption that we cannot understand a creature’s thought unless we can share it and express it as one of our own. However, this view involves an egregious equivocation on the notion of understanding, which can mean *interpreting* a meaningful representation (*verstehen*), or *explaining* it as in science (*erklären*). Failure to heed this distinction has bedevilled a long dispute in the social sciences (Slezak, 1990) and is at the heart of the vexed debates in cognitive science too (Slezak, 2002). The undoubted fact that we are unable to interpret a dog’s thoughts and express them in our own language has no bearing on the possibility of understanding them through science. Likewise, the question of semantics or meaning of mental representations in cognitive science is standardly confused between whether representations are *intelligible* and whether they are *explainable*. Searle’s (1980) notorious Chinese Room conundrum trades directly on this confusion by asking whether a fully intelligent language understander can interpret computational symbols which are the substrate of thought. Their unintelligibility to such a homunculus intruding into the system has curiously been taken as relevant to the theoretical, explanatory issue of the intentionality of symbols. However, this explanatory issue must be the quite different question of how the symbols relate to one another, the organism’s behaviour and the outside world.

The relevance of Searle’s seemingly unrelated problem to our present interests is seen from the fact that his requirement for the intelligibility of the computational symbols is essentially the requirement that the language of thought be English. Searle contrasts the meaningless computational symbols of mentalese (analogous to Chinese) which he cannot understand with those (in English) that he can understand. The much-discussed scenario of the Chinese Room is best understood, then, not as a challenge to ‘Strong AI’ but as a *reductio ad absurdum* of the conception of mental representations as intelligible to an external observer. It must be said that this conception is not Searle’s mistake alone but widely shared throughout AI and

---

<sup>3</sup> Davidson’s point seems closely related to the argument of Winch (1957) who argued that in the case of meaningful human social behaviour, the scientist is not at liberty to frame explanatory concepts at will regardless of the agent’s own conceptions.

cognitive science (Slezak, 2002). Indeed, it is evidently at the heart of Newell and Simon's (1976) 'physical symbol system' hypothesis as Newell (1986) clearly reveals:

The idea is that there is a class of systems which manipulate symbols, and the definition of these systems is what's behind the programs in AI. The argument is very simple. We see humans using symbols all the time. They use symbol systems like books, they use fish as a symbol for Christianity, so there is a whole range of symbolic activity, and that clearly appears to be essential to the exercise of mind. (1986, p. 33)

This account is noteworthy for the explicitness with which it embraces paradigmatically external symbols as models for internal representations. In AI, Nilsson (1991), too, is explicit in embracing a logicist view of external, interpreted symbols as appropriate for modelling intelligence. This conception has been embodied in AI systems such as the CYC program of Lenat and Feigenbaum, aptly caricatured by Smith (1991) as the 'Electric Encyclopedia'. Cummins (1996), too, has clearly pointed to the mistake of construing internal representations as if they are exploited by being understood. It is, of course, no accident that pictorial theories of imagery work in exactly the same way by positing representations which are intelligible to an intelligent perceiver (Slezak, 1992, 1995). This creates the illusion of an explanatory theory in each case, whereas it is the theorist who is tacitly doing the work. These reflections suggest that, notwithstanding the repudiation of a thinking-in-language account, even the orthodox computational theory of mentales may be committed to ascribing the properties of natural language to the internal code.

## 7. Belief and the concept of belief

Evidence of the ubiquitous conflation of thinking, on one hand, and the *reflective thinking about thinking*, on the other hand, is seen in Davidson's claim that a creature cannot have a belief if it does not have the *concept* of belief. This is supposedly because "Someone cannot have a belief unless he understands the possibility of being mistaken, and this requires grasping the contrast between truth and error—true belief and false belief" (1975, p. 22). Carruthers makes a similar point saying conscious thought is a matter of having second-order thoughts about other thoughts (1996, p. 157). Carruthers, too, suggests that discrimination between experiences requires "that the subject should be capable of thinking about (and so conceptualizing) its own experiences" (1996, p. 157). Nevertheless, if this is not a mere stipulation about how to use the words 'belief' and 'thinking', it seems exceedingly unsatisfactory from a theoretical point of view to deny that animals might have beliefs because they are unable to know that they have them and reflect on their truth value. It is entirely unclear why a creature could not *have* both true and false beliefs without also realizing which was the case in a particular instance. A cat can surely be correct in thinking that a mouse is in a certain hole without having the concepts of belief and truth.



As Fodor has aptly put the point:

Surely what matters to whether it's all right for me to step on the cat's tail is primarily whether it hurts him, not what he thinks about it: still less whether he thinks about what he thinks about it. (1999, pp. 12–13)

Consistently with the mistake we have noted, Davidson requires that judgements which can only be made from the external perspective of the theorist be ascribed to an agent. Just as noted in the case of pictorial representation, too, the attribution of pictorial resemblance requires a comparison and, therefore, external judgement.

Searle (1994) has also questioned Davidson's demand for second-order notions such as the *concept* of belief in the attribution of belief, but it is noteworthy that his way of expressing his contrary view subtly encourages the very mistake he attributes to Davidson. Thus, he says that animals need not have metalinguistic beliefs or concepts in order to have true or false beliefs, but he expresses his point by saying that an animal is able "to distinguish the state of affairs in which their belief is satisfied from the state of affairs in which it is not satisfied" (1994, p. 212) or an animal "can tell whether its belief is true or false" (1994, p. 212). However, such locutions are ambiguous in continuing to suggest that an animal might have beliefs about their beliefs whereas they merely *have* them. That is, strictly speaking, an animal does not "tell true from false beliefs" as Searle puts it, but only *has* true or false beliefs. *We*, as theorists, may "tell" true from false beliefs since this is a meta-linguistic judgement, and is not the same as simply *having* beliefs which may be true or false.

## 8. Ryle on thinking "in" language

Before turning directly to Carruthers' (1996) defence of the cognitive conception of language, it is instructive to note an important but neglected discussion by Gilbert Ryle (1968). Ryle pointed out that one sense in which we may be said to think in language is that in which we actually rehearse *sotto voce* words or sentences which we might or might not utter—talking to one's self, as it were, in a silent monologue. This is, in fact, the sense of 'thinking in language' which Carruthers defends as rival to Fodor's 'language of thought' account. It is undeniably a familiar and central version of the doctrine and, accordingly, its exact meaning and status must be examined carefully.

This claim concerning the substrate or vehicle of thought is defended by Carruthers, and it is the very intelligibility of such a claim which I will be concerned with. In this concern I am, indeed, merely defending Ryle's own view which attacks the presupposition that "when thinking occurs, there occur, internally or externally, things or symbols that the thinker thinks in" (1968, p. 12). Ryle says:

I want to attack this presupposition. I want to deny that it even makes sense to ask, in the general case, what special sort or sorts of things we think *in*. The very collocation of 'think' with 'in so and so' seems to me factitious ... (1968, p. 13)

Ryle's skepticism seems to be ruling out, not only the sort of claim that we think in a natural language such as English, but also Fodor's computational 'language of thought' hypothesis. Ryle does not go much beyond raising the skeptical question, but it is significant that he mentions, *en passant*, among the equally problematical cases are those in which we claim to see things in our mind's eye, taking this to involve mental pictures of some kind. Of course, the notorious 'imagery debate' (Block, 1981) centred on just such a supposition. I will show that this controversy has significant analogies with the question of language and thought of interest here. Ryle's passing remark signals an insight into the parallel difficulties between claims that we think in language and claims that we visualize in pictures. Ironically, Ryle's passing comparison between these problems is unwittingly, confirmed by Carruthers who explicitly models his own theory on Kosslyn's pictorial theory of visual imagery (1998, p. 97). It is telling that Carruthers fails to cite Ryle's article. Of course, the wisdom of Ryle's prescient warnings have been amply supported by the trenchant critiques of pictorial theories by Pylyshyn (1973). Consequently, at the very least, Carruthers cannot simply assume that pictorial accounts of imagery provide an unproblematic model for his own theory of language and thought. On the contrary, the uncritical assimilation of the pictorial model to his own brings into relief its shortcomings.

However different the cognitive domains, the common question concerning vehicles of thought presents common pitfalls. In Ryle's view, just as visual imagery must be explained without pictures in the head, so thought must be explained without sentences in the head. However, Ryle's concerns do not necessarily apply to mentalese. His skepticism concerning the very idea that there is something we can be said to think 'in' are confined to a particular conception of these objects and what exactly it means to think "in" them. Appropriately conceived, mentalese cannot be described as something we think "in" as Carruthers and others construe this notion. As Dennett (1978, p. 101) aptly put it, an appropriate theory would have ideas to "think for themselves".

Carruthers (1996) sees his account as a rival to Fodor's mentalese theory, but this assumes that his 'cognitive conception' of language can be coherently formulated as an explanatory empirical hypothesis and, further, that mentalese is something we might think "in" in the same sense. However, I will suggest that Carruthers' strenuous defence of the doctrine falls into Ryle's trap and is an extended attempt to make a case for precisely the idea that Ryle describes as "factitious".

## 9. Precursors and persistence

Ryle's warnings of the notorious trap have an earlier provenance: In a recent book, Yolton (1996, p. 43) mentions the anonymous author of a pamphlet written in 1705 titled *Philosophick Essay Concerning Ideas* who wrote:

... in considering the *Mind*, some men do not sufficiently abstract their Thoughts from *Matter*, but make use of such Terms as can properly relate to *Matter* only, and apply them to the *Mind* in the same Sense as they are spoken

of *Matter*, such as *Images* and *Signatures, Marks, and Impressions, Characters* and *Notes of Things*, and *Seeds of Thoughts* and *Knowledge*. (quoted in Yolton 1956, p. 96)

These remarks point to the error of taking external objects or ‘Matter’ such as ‘Images’, ‘Signatures’ or ‘Characters’ as the model for our internal, mental entities. This is precisely to warn of taking external communicative symbols such as those of natural language as possible vehicles of thought. Yolton is concerned to trace the long history of philosophical disputation concerning ‘ideas’ as the immediate objects of thought and, thereby, a topic which may seem remote from our present concern with language and thought. However, my contention here is that the two problems are intimately related, as Ryle’s allusion to visual imagery clearly suggests. The ‘cognitive conception’ of language takes words—paradigmatically external representations—as the internal objects of thought in exactly the same sense that a long tradition has taken ‘ideas’—namely, as inner mental objects which are the vehicles of representation. The notorious problem of ideas arose from taking external representations such as words or pictures as appropriate models for inner, mental representations. Significantly, Antoine Arnauld (1683) too, following Descartes, warned against those who “believe that they will find some analogy in the body which will allow them to understand how we see with the mind everything that we conceive, and above all material things” saying further “Only confusion can result from trying to explain the one by the other, and one of the most pervasive sources of error is our constant application of the properties of the mind to bodies and properties of bodies to the mind” (1683/1990, p. 60). Arnauld’s target was Malebranche’s doctrine of representative ideas seen later also in Locke, and in one form or another still pervasive in contemporary cognitive science, most notably in the imagery debate (Slezak, 2002, in press).

Cummins (1996) has recently said “*the* topic in the philosophy of mind for some time now has been to explain the relevant notion of mental representation ...” (1996, p. 1). This is something of an understatement, since the same problem was central to the dispute between Arnauld and Malebranche 400 years ago. In its current version, the issue of intrinsic versus derived intentionality is precisely the problem of explaining how inner mental symbols might obtain their semantic content or meaning and how this differs fundamentally from external representations such as words or pictures. The idea that we might think in natural language is evidently meant to “solve” this problem by assimilating internal, original intentionality with external, derived intentionality.

## 10. Carruthers: Conscious thought and introspection

Carruthers’ defence of the doctrine that we think in language is important in being a vigorous defence of a discredited idea, and also because it explicitly defends introspection as its source. Carruthers sees the usual scientific evidence cited against thinking in language as equivocal saying “it is too early, as yet, for any firm

conclusions to be drawn from the available scientific data” (1996, p. 49). Accordingly he appeals to what he calls “a very different set of data, namely that delivered by ordinary introspection” (1996, p. 49). Such data of introspection have obvious connections with his central claim that it is only *conscious* thought which may be claimed as being “in” natural language. Of course, this is precisely the case in which we silently talk to ourselves or rehearse sentences of a language. For Carruthers, conscious thought is identified not with the kind of ordinary awareness which animals might have, but with a second order, reflective thinking about thinking. As Fodor’s (1999) parody makes clear, Carruthers is concerned not with thought *per se* but with thinking about thought (1996, p. 221). Thus he explains that he is concerned with “a faculty of higher-order thought, in which thinkings are themselves sometimes available to further thought” (1996, p. 157). That is, for Carruthers, conscious thoughts “constitutively involve” natural language in the sense that they are *about* natural language. In this sense, language functions as the object or content of thought. Our conscious thinking is “in” natural language in the sense that

... we mostly think (when our thinking is conscious) by imaging sentences of natural language, and trains of thought consist of manipulations and sequences of such images. We have access to the forms of our thoughts because a record of each imaged sentence is briefly held in short-term memory, so that we can recall what we have just imaged. ... my image is immediately and non-inferentially imbued with content, just as if I had heard that sentence uttered aloud. (1996, p. 228, 229)

Of course, this is precisely the pictorial model of seeing with the mind’s eye transposed to a different modality of imagination. Remarkably, mere suspicions about this parallel between Carruthers (1996) earlier account and the pictorial theory are explicitly confirmed by Carruthers’ (1998) more recent discussion. However, in thinking that he might exploit the parallel to defend his theory Carruthers only provides the strongest confirmation of the aptness of the independent criticisms which parallel those against pictorialism. The parallels are instructive through revealing independent ways of making the same seductive mistake in reasoning about the mind.

By endorsing the parallel between the two models, Carruthers helps us to understand exactly what he has in mind concerning the way in which we think “in” language. Notoriously, Kosslyn’s pictorial account of visual imagery postulates a short-term, visual memory buffer which serves as an alternative source of input stimuli to the apparatus of higher vision. The uninterpreted display on this visual buffer “is processed by the higher visual system in the normal way to yield a quasi-visual percept” (1998, p. 97). In this way the pictorial or iconic images in the buffer serve as surrogate objects for the visual system and imagery is thereby supposed to afford the possibility of re-inspection and re-interpretation of these internal representations.<sup>4</sup>

<sup>4</sup> Experimental tests of the divergent predictions of Pylyshyn’s tacit knowledge theory and Kosslyn’s pictorial theory have not supported the pictorial account concerning reinterpretation of images (Slezak, 1992, 1995).

Carruthers indicates the parallel with his own account which is complete to the extent of positing sentential images in a short-term memory buffer just like Kosslyn's visual buffer. Just as we seem to see pictures, so Carruthers says "we seem to hear the meaning of the utterance" (1996, p. 229) which he says is "just as if I had heard that sentence uttered aloud" (1996, p. 229). Thus, "we mostly think (when thinking is conscious) by imaging sentences of natural language" (1996, p. 228).

Of course, as Descartes clearly explained in his *Dioptrique*, such an account invokes an intelligent homunculus by positing an uninterpreted image which requires the full visual system to perceive it. Carruthers' account substitutes a natural language sentence for the pictorial image, and substitutes central linguistic, understanding processes for higher vision. That is, just as higher visual systems perceives the image, so higher language mechanisms understand an inner sentence. By analogy with visual imagery,

... people would be able to generate sentences of inner speech in imagination, which would get interpreted by the language system in the normal way, and whose contents would be available to meta-representational thought. (1998, p. 99)

At the very least, in view of the notoriously controversial status of pictorialism, Carruthers cannot uncritically co-opt it to bolster his own theory. It is precisely in those respects in which his theory is analogous to pictorialism that Carruthers' theory is problematic. Far from providing independent support, the analogy with Kosslyn's theory reveals the fatal shortcomings of his theory.

## 11. Thinking in mentalese?

The imagery model makes it clear why Carruthers thinks that his theory is a challenge to Fodor's LOT account:

Even if we do possess some sort of Mentalese, we have no access, surely, to its forms and structures—we are incapable of thinking a thought in Mentalese, and then thinking (non-inferentially) about the Mentalese sentence we have just entertained. So without a public language we should be incapable of conscious (propositional) thinking ... (1996, p. 256)

That is, mentalese is not the sort of thing that we can think about. Of course, Carruthers is correct in this, but it counts against his own view rather than Fodor's. Carruthers' entire case rests on the undoubted fact that we are sometimes consciously aware of sentences in natural language, that is, "the introspective datum that much of our conscious thinking appears to take place in natural language" (1996, p. 228). However, the fact that we are never conscious in this way of mentalese does not count against mentalese as a posit but provides rather a *reductio ad absurdum* of the account which requires it.

Carruthers' entire case is carefully restricted to the case where it seems intuitively most plausible—namely, conscious, introspectible, silent 'talking to ourselves'. However, this plausibility is only the force of tautology: The claim becomes merely that the language that we think *about* is the language we speak; that is, when we are conscious of silently talking to ourselves it is in our own natural language—hardly a startling or theoretically interesting conclusion. Certainly such avowals of the contents of introspection are not denied by any theory of mentalese. Moreover, on Carruthers' own account, as one among the many possible objects of our thinking, it is difficult to see how natural language could possibly present any threat to Fodor's mentalese as the substrate of all such thought. It is no accident that Pylyshyn's charges against pictorial theories of imagery included the complaint that they confused properties belonging to the *objects* of imagery with those of the imagery mechanism itself. That is, just as images representing space do not require the representations to actually *have* spatial properties, so imagining speech does not require the representations to actually be linguistic.

Although presenting his account as a rival to Fodor's, Carruthers concedes that much of our thinking may, in fact, be conducted in Fodor's language of thought and, therefore, he wishes only to claim that *some* of our thinking is in natural language (1996, p. 252). Since Fodor's account purports to be universal, Carruthers claims that it is mistaken at least in one important domain of our mental lives, namely, conscious propositional thought. In this way, Carruthers can also concede all the empirical evidence standardly cited concerning animals, pre-linguistic children and thoughts which are not plausibly linguistic in any sense. He restricts the scope of his claim about thinking in language to *human conscious thinking*.

In the foregoing quotation Carruthers reveals his fundamental misunderstanding of Fodor's mentalese. Mentalese is not incompatible with conscious propositional thinking because it is not the sort of symbolic system to which we have conscious access, nor is such access required in order for conscious propositional thinking to take place. Fodor's mentalese is at the level Pylyshyn (1984) calls the level of cognitive architecture and which forms the substrate or symbol system in which conscious and other thoughts are encoded. The symbols are not themselves the object or contents of these thoughts. Carruthers' thesis that we think *about* language in some cases when we introspect is uncontroversial but can hardly count as a competitor or counter-example to Fodor's account. Thus, Carruthers attempts to establish his account on a false dichotomy in saying:

The main focus for debate will concern *which* sentences are constitutive of our (conscious propositional) thoughts—those of Mentalese, or those of natural language. (1996, p. 39).

If my conscious thinkings take place in English, as I have claimed on the basis of introspection that they do, then of course they do not take place in any other natural language, let alone in a supposedly innate, universal, symbolic system such as Mentalese. So if the argument from introspection is allowed to stand, ... Fodor and others are wrong to claim that all thoughts are expressed in Mentalese. (1996, p. 55)

Carruthers is trading on an ambiguity in our idiomatic expression ‘thinking in language’ and relies on the equivocation to challenge Fodor’s theory. Mentalese does not purport to be a language we think in, where this is understood as the language we are consciously aware of rehearsing when we talk to ourselves. As Ryle noted, the latter sense of “thinking in” a language does not establish any claim about the medium or vehicle of thought.

## 12. Introspection systematically misleading?

Carruthers explicitly acknowledges the possibility that introspection might be misleading in some systematic way and he recognizes that this would defeat his case against Fodor’s LOT account. However, he does not contemplate the most relevant and serious kind of challenge to introspection, instead drawing a mistaken conclusion from Fodor’s mentalese. Thus, in his exposition of Fodor’s general account, Carruthers observes that, if correct, Fodor’s picture “forces us to give up an important ingredient of common-sense belief—namely, that we often think aloud *in* natural language” (1996, p. 88). Accordingly, he says, “This introspective data will need to be explained, or explained away, by anyone who wishes to insist that all thinking takes place in Mentalese” (1996, p. 5). However, as already suggested, Carruthers fundamentally misconceives the issues at stake here. He sees the vindication of Fodor’s mentalese as denying “the claim that we entertain our conscious thoughts in natural language” (1996, p. 135). However, the truth of Fodor’s theory has no such implications for the contents of conscious thought. It is undeniable that we entertain some of our conscious thoughts in natural language, the only question being how this is to be explained. Thus, we are not misled about whether we *seem* to be thinking in language and that we are silently talking to ourselves. Introspection cannot be misleading about this fact of subjective awareness, but only about the quite different claim that the objective underlying process itself is in natural language. The latter claim is quite different and the evidence relevant to it is quite different from introspective appearances.

## 13. Theories of no imagery, no inner speech and no consciousness?

It is striking that in the independent debates concerning visual imagery (Kosslyn, 1980) and consciousness (Searle, 1992) there are analogous charges that one side is simply ignoring the most obvious, uncontroversial facts of inner experience. Kosslyn (1980, p. 30) has revealingly described Pylyshyn’s (1973) ‘tacit knowledge’ alternative to the pictorial theory as a “no imagery” account as if it purports to deny the reality of visual imagery as such. This is an egregious mistake, but it is most revealing about the thinking which led to it and the position in whose service it is deployed. Similarly, it is perhaps no coincidence that a parallel accusation has been made by Searle (1992, p. xi) who charges that materialist philosophers “routinely denied what I thought were simple and obvious truths about the mind”—namely,

that “We all have inner subjective qualitative states of consciousness...” It is no accident that Carruthers also defends his account against what he takes to be an implication of mentalese, namely, that it denies the manifest evidence of introspection. This is the same mistake in each case since it is not the evidence of introspection as such being denied but rather a particular explanation of it.

The mistake in each case is to conflate introspection with a theory of it, taking introspection as itself somehow explanatory rather than merely constituting data to be explained. Thus, Carruthers himself gives ample warrant for undermining his own theory by conceding that it would be defeated if introspection could be shown to be systematically misleading. We may fairly conclude that, on his own account, his theory is falsified—not because the introspections themselves are mistaken as phenomenology, but because they are systematically misleading about the vehicles of thought and the underlying mechanisms.

#### **14. Semantics: how things reach out and touch someone?**

There is, after all, a substantive, scientific question concerning the relation of language to thought which, as Carruthers notes (1996, p. 4), concerns the semantics of natural language.

If some version of the communicative conception of language is correct, then it must be possible to provide a semantics for the expressions of natural language in terms of prior notions of thought and intention. It must also be possible to provide a semantics for thought, in turn, without reintroducing natural language into the story. (1996, p. 4)

This is, of course, just the recalcitrant problem of intentionality or ‘psycho-semantics’ whose burgeoning literature attests to the fact that there is a consensus, at least, on Fodor’s judgement that “of the semanticity of mental representations we have, as things now stand, no adequate account” (1990, p. 28). Typically, Stalnaker, too, says “There is little agreement about how to do semantics, or even about the questions that define the subject of semantics” (1991, p. 229). Likewise, Smith confesses “It should be admitted that how this all works—how symbols ‘reach out and touch someone’—remains an almost total mystery” (1987, p. 215)

Carruthers recognizes that any account of the language of thought, even if it is the same as the language of speech, will have to face these notoriously intractable difficulties of semantics. Accordingly, he surveys the central issues as they have been most prominently formulated over two decades by Jerry Fodor (1981, 1987, 1990, 1994). Carruthers’ treatment of these issues is of interest because they provide further ample evidence supporting the diagnoses I have offered.

After rehearsing familiar problems concerning the semantics of mentalese such as that of misrepresentation (Dretske, 1994; Fodor, 1994), Carruthers offers a response to Fodor’s puzzlement about this issue. The details of the problem are not relevant



here except to the extent that they illuminate Carruthers' account. Fodor noted a puzzle arising for causal theories of mental content since causal co-variance seems unable to explain misrepresentation. On a co-variance view, the mentalese term 'mouse' means mouse because a tokening of 'mouse' is generally caused by the presence of a mouse. That is, only mice cause tokenings of 'mouse'. However, since other things may occasionally be mistaken for mice, a shrew might cause the tokening of 'mouse'. But on the causal theory this cannot be a misrepresentation because whatever causes the tokening of a term is ipso facto what it means. Fodor considers the possibility that, on the causal theory, then, the tokened term may mean 'shrew-or-mouse'. As Fodor recognizes, this solution is not satisfactory because the relation of a mentalese term 'mouse' to mice is not the same as the relation of 'mouse' to shrews or any other non-mice. Specifically, there is an asymmetric causal dependence in which shrews could not have caused tokenings of 'mouse' unless mice could, whereas mice could cause tokenings of 'mouse' even if shrews couldn't. The details are of no further interest here since they are unnecessary for appreciating the way in which Carruthers construes the issues at stake. His response to problems raised in Fodor's account is addressed as follows:

The intuitive worry is this: how plausible is it that our judgements of representation, or of misrepresentation, are actually guided by *thoughts about* asymmetric causal dependence? When someone says 'Mouse' in the presence of a shrew . . . how plausible is it that my judgements about what they mean will be driven by considerations of causal asymmetry? Certainly these do not appear to enter into my conscious practice at all. (1996, p. 94; emphasis added.)

First, we see here a crucial, though mistaken, reliance on *actual utterance* in natural language as when someone *says* 'Mouse'. Clearly, the central issue for Fodor's psychosemantics is not speech but thought which need not be accompanied by the verbal utterance of any term at all. For Fodor, the issue of semantics concerns the meaning of terms of mentalese, not spoken words which might express them. By collapsing the two cases, Carruthers is legislating his theory of thinking-in-language correct and begging the central question at issue.

Second, and more serious, is Carruthers' gross misunderstanding of the status of Fodor's asymmetric causal dependence. Fodor offers this as part of his philosophical analysis *qua* theorist and not *qua* speaker. That is, the notion of asymmetric causal dependence is an explanatory concept which might help make sense of the theoretical puzzles about mentalese, and it is definitely not referring to a concept that a speaker has access to or might possibly apply as part of any "judgements" or consideration about the use of language. Carruthers clearly takes Fodor to intend the concept as something which might "enter into my conscious practice" and be part of "our judgements of representation" as *speakers*. The idea that we might be "actually guided by thoughts about asymmetric causal dependence" or "driven by considerations" of this kind is an absurdity which cannot be attributed to Fodor's analysis. Carruthers is correct in seeing that we do not employ such notions in our judgements, but this is utterly irrelevant to the issues at stake in Fodor's analysis.

We see a clue to the problem in Carruthers' slide between two distinct applications of crucial notions of meaning in a phrase such as "by 'mouse' this person may mean..." or "I can use 'mouse' to mean *mouse*" (1996, p. 91). These expressions are ambiguous between a sense in which I consciously intend to use a term, and a sense in which I "use" a term at a sub-personal level as part of consciously inaccessible processes. The term "use" itself suffers the same ambiguity. That is, I might explain my use of the term "bank" by saying that I mean "financial institution" and not "river edge". However, this is importantly different from the case in which my mentalese terms might "mean" something and how I might be said to "use" them. Failure to observe this distinction is catastrophic for theorising in this domain. Carruthers' equivocation is no mere slip, but essential for his central thesis which depends crucially on a sense of the expression in which "I use a term to mean something" in a deliberate act of awareness. This sense of the crucial words cannot be applied to the terms of mentalese which are not "used" or "meant" in this sense by a person at all. In this sense, I don't literally use my terms of mentalese. The same ambiguity infects the term 'represent' which also permits a sense in which I use something in order to represent something else. As Dennett (1978) and Cummins (1996) have pointed out, this cannot be the sense in which mental representations do their work. In Carruthers' dependence on these expressions to denote self-conscious, personal level facts, we see clearly the same conception noted earlier—namely, a notion of language and thought which is exclusively that of consciously accessible, introspectible experience.

Thus, leaving no doubt about my attributions, regarding the dilemma about whether a term means 'mouse' or 'shrew-or-mouse' Carruthers asks about "what steps I would take to resolve the issue?" His answer is revealing: He says "I would *not* ask whether their tokens of MOUSE would still have been caused by shrews like that one, if tokens of it had not also been caused by mice; nor conversely etc. ..." (1996, p. 94). Again, this is a clear appeal to what the person would or would not say about their use of a term. His reason for not asking the person his question is not because asking is irrelevant, but because Carruthers (rightly) recognizes that the person would not know the answer. However, Carruthers takes this to prove Fodor's speculation wrong whereas the person could not know the answer even if Fodor were right! Asking "What do you mean by X" is not the appropriate solution to Fodor's problem since neither the meaning of mentalese terms nor their putative asymmetric causal dependence are available in this way. The answer to Carruthers' question about how to resolve the issue must involve the philosophers' theoretical dilemma and not the subject's.

## 15. Conclusion: the phenomenological fallacy

Carruthers writes:

Remember that according to the cognitive conception, when a speaker utters a sentence their utterance will express a thought by *constituting* it, not by encoding or signaling it. (1996, p. 122)

How sentences of a natural language could conceivably “constitute” thought is utterly unclear because such sentences require interpretation. It is an illusion that the sentences themselves might function in cognition because it is we as theorist who are surreptitiously doing the work by using our own intelligence to compensate for the deficiencies of the theory. Pictures also appear to serve imagery by somehow “constituting it”, but again only because we as *theorists* can understand or perceive them. Thus, Kosslyn (1994, p. 14) claims that pictorialism is vindicated by the discovery of retinotopic maps of distal stimuli on the brain of experimental animals. Of course, these are pictures, but not for the animal, only the theorist.

I have suggested that the very naturalness or indeed compellingness of the view we have been considering is, paradoxically, the strongest evidence against it. This irony is understood when it is seen that the compellingness of the view does not derive from its systematic scientific merits but from the seductive persuasiveness of a certain misleading picture which foists itself upon us as a consequence of certain inherent cognitive arrangements. The very possibility of introspective self-reflection encourages the mistake of thinking that the “objects” of our internal experiences are similar to their external vehicles. This *ad hominem* charge gains considerable force from the surprising ubiquity of the mistake across seemingly independent domains and times. Thus, to take just one example, it is remarkable that in a well-known article by Place (1956), first to articulate a materialist theory of mind, he writes under a heading ‘The Physiological Explanation of Introspection and the Phenomenological Fallacy’:

... for all its emotional appeal [Sherrington’s argument for dualism] depends on a fairly simple logical mistake, which is unfortunately all too frequently made by psychologists and physiologists and not infrequently in the past by the philosophers themselves. This logical mistake, which I shall refer to as the ‘phenomenological fallacy’, is the mistake of supposing that when the subject describes his experience, when he describes how things look, sound, smell, taste or feel to him, he is describing the literal properties of objects and events on a peculiar sort of internal cinema or television screen. . . (1956, p. 37)

Significantly, Place observes in a footnote that his case “often seemed a lost cause” and it seems likely that he is reflecting on the persistent difficulty of arguing against the strong contrary intuitions derived from introspection.

I have argued that if imagery cannot plausibly be in pictures, then thoughts cannot be in language either, and for the same reasons. Once we distinguish the idea of a representation and its meaning *qua* theoretical posit from representation and meaning *qua* object of apprehension, it appears much less clear what properties we may justifiably ascribe to the theoretical posit—that is, if we are to avoid the mistake noted by Pylyshyn of attributing properties of the world (our external representations) to our internal, mental states. *Pace* Carruthers, it is not to deny that mental life may *appear* to be a silent soliloquy to propose a non-linguistic explanation of this introspective fact. In this way, an *ad hominem* strategy significantly different from Magee’s may be a cogent argument after all: It discredits the argument not by

challenging the rival introspections themselves, but by undermining the source or warrant of their persuasiveness.

The concerns I have been raising are hardly new. A depressing feature of contemporary philosophy and cognitive science is the extent to which they have lost sight of their past and the consequent recycling of debates which are essentially unchanged since the seventeenth century. In the present case, we need go back no further than Ryle who said “I want to deny that it even makes sense to ask, in the general case, what special sort or sorts of things we think in.”

## References

- Armstrong, D.M., 1980. Epistemological Foundations for a Materialist Theory of the Mind. In: *The Nature of Mind and Other Essays*, St. Lucia: University of Queensland Press, pp. 32–54. Originally in *Philosophy of Science*, 1973.
- Arnauld, A., 1683/1990. *On True and False Ideas*. Translated with Introductory Essay by Gaukroger, S., Manchester University Press.
- Block, N., 1981. *Imagery*. MIT Press, Cambridge, MA.
- Carruthers, P., 1996. *Language, Thought and Consciousness: An Essay in Philosophical Psychology*. Cambridge University Press, Cambridge.
- Carruthers, P., 1998. Thinking in language? Evolution and a modularist possibility. In: Carruthers, P., Boucher, J. (Eds.), *Language and Thought: Interdisciplinary Themes*. Cambridge University Press, Cambridge.
- Carruthers, P., Boucher, J., 1998. *Language and Thought: Interdisciplinary Themes*. Cambridge University Press, Cambridge.
- Cummins, R., 1996. *Representations, Targets and Attitudes*. Bradford/MIT Press, Cambridge, MA.
- Davidson, D., 1975. Thought and talk. In: Guttenplan, S. (Ed.), *Mind and Language*. Clarendon Press, Oxford, pp. 7–23.
- Davidson, D., 1997. Seeing through language. In: Preston, J. (Ed.), *Thought and Language*. Cambridge University Press, Cambridge, pp. 15–27.
- Dennett, D., 1978. A cure for the common code. In: *his Brainstorms*. Vermont: Bradford Books, pp. 90–108.
- Devitt, M., 1984. Thoughts and their ascription. In: French, P., Uehling, T., Wettstein, H. (Eds.), *Midwest Studies in Philosophy IX*. University of Notre Dame Press, Notre Dame, pp. 385–420.
- Devitt, M., Sterelny, K., 1987. *Language and Reality*. Basil Blackwell, Oxford.
- Dretske, F., 1994. Misrepresentation. In: Stich, S., Warfield, T.A. (Eds.), *Mental Representation*. Blackwell, Oxford, pp. 157–173.
- Duhem, P., 1906/1962. *The Aim and Structure of Physical Theory*. Princeton: Princeton University Press.
- Fodor, J.A., 1975. *The Language of Thought*. Crowell, New York.
- Fodor, J.A., 1981. *Representations*. MIT Press, Cambridge, MA.
- Fodor, J.A., 1987. *Psychosemantics*. MIT Press, Cambridge, MA.
- Fodor, J.A., 1990. *A Theory of Content*. MIT Press, Cambridge, MA.
- Fodor, J.A., 1994. *The Elm and the Expert*. MIT Press, Cambridge, MA.
- Fodor, J.A., 1998. Do we think in mentalese? In *Critical Condition*. MIT Press, Cambridge, MA.
- Fodor, J.A., 1999. Not so Clever Hans. *London Review of Books*, February, 12–13.
- Glock, H., 1997. Philosophy, thought and language. In: Preston, J. (Ed.), *Thought and Language*. Cambridge University Press, Cambridge, pp. 151–169.
- Harman, G., 1973. *Thought*. Princeton University Press, Princeton.
- Jackendoff, R., 1997. *The Architecture of the Language Faculty*. MIT Press, Cambridge, MA.
- Kosslyn, S., 1980. *Image and Mind*. Harvard University Press, Cambridge, MA.
- Kosslyn, S., 1994. *Image and Brain*. MIT Press, Cambridge, MA.

- Magee, B., 1997. *Confessions of a Philosopher*. Weidenfeld & Nicolson, London.
- McDowell, J., 1994. *Mind and World*. Harvard University Press, Cambridge, MA.
- McGinn, C., 1997. *The Character of Mind*. Oxford University Press, Oxford.
- Moore, G.E., 1942. *Autobiography*. In: Schilpp, P.A. (Ed.), *The Philosophy of G.E. Moore*. Northwestern University Press, Evanston, pp. 3–39.
- Newell, A., 1986. The Symbol Level and the Knowledge Level. In: Pylyshyn, Z. and Demopoulos, W. (Eds.) *Meaning and Cognitive Structure*. Ablex, Norwood, NJ, pp. 31–39.
- Newell, A., Simon, H.A., 1976. Computer science as empirical inquiry: symbols and search. *Communications of the ACM* 19, 113–126.
- Nilsson, N.J., 1991. Logic and artificial intelligence. *Artificial Intelligence* 47, 31–56.
- Pearson, D., Reisberg, D. and Kosslyn, S., 2000. *Imagery Survey*. Email communication.
- Pinker, S., 1994. *The Language Instinct*. Allen Lane Penguin, London.
- Place, U.T., 1956. Is consciousness a brain process? *British Journal of Psychology*; reprinted in Beakley, B. and Ludlow, P. (Eds.) *The Philosophy of Mind: Classical Problems/Contemporary Issues*. Cambridge, MA: Bradford/MIT Press, 33–39.
- Pylyshyn, Z., 1973. What the mind's eye tells the mind's brain: a critique of mental imagery. *Psychological Bulletin* 80 (1), 1–24.
- Pylyshyn, Z., 1984. *Computation and Cognition*. MIT Press, Cambridge, MA.
- Ryle, G., 1968. A puzzling element in the notion of thinking. In: Strawson, P.F. (Ed.), *Studies in the Philosophy of Thought and Action*. Oxford University Press, Oxford, pp. 7–23.
- Searle, J., 1980. *Minds, brains and programs*. *Behavioral and Brain Sciences* 3, 417–424.
- Searle, J., 1992. *The Rediscovery of the Mind*. MIT Press, Cambridge, MA.
- Searle, J., 1994. *Animal Minds*. In: French, P.A., Uehling, T.E., Wettstein, H.K. (Eds.), *Philosophical Naturalism: Midwest Studies in Philosophy*, Vol. XIX. University of Notre Dame Press, Indiana.
- Slezak, P., 1990. Man not a subject for science? *Social Epistemology* 4 (4), 327–342.
- Slezak, P., 1992. When can images be reinterpreted: non-chronometric tests of pictorialism. In: Kruschke, J.K. (Ed.), *Proceedings of 14th Conference of the Society for Cognitive Science*. Hillsdale, NJ: Lawrence Erlbaum, 124–129.
- Slezak, P., 1995. The 'philosophical' case against visual imagery. In: Slezak, P., Caelli, T. and Clark, R. (Eds.) *Perspectives on Cognitive Science: Theories, Experiments and Foundations*. Ablex, Norwood, NJ, pp. 237–271.
- Slezak, P., 1999. Situated cognition: empirical issue, paradigm shift or conceptual confusion? In: Wiles, J. and Dartnall, T. (Eds.) *Perspectives on Cognitive Science*, Vol. 2. Ablex, Stamford, CT, pp. 69–98.
- Slezak, P., 2002. The Tri-partite model of representation. *Philosophical Psychology*, in press.
- Smith, B.C., 1987. The Correspondence Continuum. *CSLI Report* 71–87.
- Smith, B.C., 1991. The owl and the electric encyclopedia. *Artificial Intelligence* 47, 251–288.
- Stalnaker, R., 1991. How to do semantics for the language of thought. In: Loewer, B., Rey, G. (Eds.), *Meaning in mind*. Basil Blackwell, Oxford, pp. 229–237.
- Stich, S., 1992. What is a theory of representation? *Mind* 101, 402. 243–261.
- Weiskrantz, L., 1997. Thought without language: thought without awareness? In: Preston, J. (Ed.), *Thought and Language*. Cambridge University Press, Cambridge, pp. 127–150.
- Woodfield, A., 1982. *Foreword*. *Thought and object: essays on intentionality*. Clarendon Press, Oxford.
- Wilkes, K.V., 1997. Talking to cats, rats and bats. In: Preston, J. (Ed.), *Thought and Language*. Cambridge University Press, Cambridge, pp. 177–195.
- Winch, P., 1957. *The Idea of a Social Science*. Routledge, London.
- Yolton, J.W., 1996. *Perception & Reality: A History from Descartes to Kant*. Cornell University Press, Ithaca.