Thought, language and mental representation

Jonathan Trigg (jon.trigg@louisiana.edu)
Program in Philosophy, UL Lafayette
Lafayette, LA  70504 USA

Michael Kalish (kalish@louisiana.edu)
Institute of Cognitive Science, UL Lafayette
Lafayette, LA  70504 USA

Abstract
We examine the proposal that thinking is a combinatorial operation on mental representations, and argue that it cannot be. If the argument is successful it shows that cognitive science cannot explain intelligent linguistic behavior by explaining what thinking is. We point out that this does not impugn the practice of cognitive scientists interested in human language, which, properly understood, consists in the framing and testing of hypotheses about the causally necessary enabling conditions of intelligent linguistic behavior.

Keywords: Wittgenstein; thought, language, mental representation, language of thought.

Introduction
We regard understanding as the essential thing, and signs as something inessential. (Wittgenstein, 1974)

...the limits of possible thought are the limits of the possible expression of thought.
(Bennett & Hacker, 2003)

If strong Wittgensteinian cross currents still ran into the mainstream of contemporary philosophy of mind and language, these philosophical waters would be much more turbulent than they now are. Indeed it is not even clear that they would be running in roughly their present direction. As things stand the river flows wide and slow, almost undisturbed by substantial impediments to its progress, and serious attempts to change its course are liable to seem naive or over ambitious – uncomprehending of the forces at work.

The argument we present here, which is Wittgensteinian in spirit, is meant to push quite hard against the prevailing drift. It’s an argument to the conclusion that thinking is not mental representing and thoughts are not mental representations. We are not, by any means, the first to make this sort of argument. Indeed, on a plausible reading of his two great works, the very theory Wittgenstein (1921/2001) defends in the Tractatus is the theory he repudiates in the Philosophical Investigations (Wittgenstein, 1958), and at its heart is the thesis that thinking is mental representing. Our aim here is to say precisely what it is about the claim that thoughts are mental representations (and that thinking is mental representing) that runs counter to the direction in which Wittgenstein’s later arguments lead. This, we believe, is not a direction cognitive scientists should be reluctant to travel in; for it takes us further and further away from Cartesian conceptions of mentality that can only hamstring research.

Overview of the argument and some preliminary points.
The argument we present here starts with the premise that to think that something is so is to perform combinatorial operations on representations, and it moves to the negative conclusion that to think that something is so cannot be to perform combinatorial operations on mental representations. This should not be taken to imply that thinking is a combinatorial operation on non-mental representations like words and sentences. Certain cases of thinking may be that1, but there is nothing in the argument which entails the strong view that to think is to speak silently to oneself in a natural language. The aim is only to rule out a widely held view about what thinking is, not to defend a competing one. The view that is ruled out is that for Sam to judge that ducks run is for her to perform an operation on representations of a radically different kind than those on which she would operate were she to make a corresponding assertion. On such a view when she asserts that ducks run, operations are performed on two sorts of representation, one mental one not; and when she judges that ducks run but asserts nothing they are performed only on mental representations. On this sort of view internal mental representing need not be connected to external spoken representing in any way; so what a creature can say sets no limits on what it can think.

If the argument of this paper is successful we will have shown that it can’t be because linguistic behaviour is linked with hidden events and processes that it counts as linguistic. This is liable to upset very widely held assumptions both about the nature of thinking and about the shape of cognitive scientific explanations. On these assumptions, what is important about Sam’s assertion that ducks run is how it stands to various items or events in, or states of, her brain. Illuminating explanations will consist in claims about this standing. That implies a certain conception of what makes assertions into meaningful speech acts and differentiates them from grunts and squawks. On it, for Sam

---

1 Bennett & Hacker (2003) contains an excellent series of reminders of how many different things we ordinarily call thinking.
to assert that ducks run is for her to make known the occurrence of a certain sort of mental act – a judgement, and for her to judge that ducks run is for certain syntactical operations to be performed (by Sam, or perhaps by a mind, a brain, or a brain-part) on certain information bearing states or structures (mental representations), which are themselves states of, or items/structures in, brains (or are realised in such states, items or structures). If that is what assertions are then linguistic behaviour must be an incidental outward accompaniment of mental representation. It is only because linguistic behaviour serves to indicate the occurrence of internal processes that it counts as linguistic (the parrot’s squawks may sound like bits of linguistic behaviour but they aren’t), but mental representations counts as mental representations whether or not they happen to be indicated by observable behaviour. On this view – which is profoundly Cartesian - the link with mental representation is an essential feature of linguistic behaviour, but the link with behaviour is an inessential feature of mental representation.

This conception of the relation between thought and linguistic behaviour could easily seem an essential feature of a cognitivist account of language.2 The aim of cognitive science is precisely to explain intelligent behaviour by appeal to internal processes, and the distinction between thinking and speech looks like a kind of paradigm of the distinction between intelligent behaviour and internal process (or processing). If we accept this appearance at face value, and if we identify mental representation with thinking, it will seem that a theory of mental representation will perfectly conform to the basic aim of cognitive science by providing an explanation of linguistic behaviour by appeal to internal goings on.

These appearances are deceptive. On our view, cognitive science can retain its commitment to explain behaviour by appeal to inner processes, events and states, whilst abandoning the Cartesian project of explaining linguistic behaviour by appeal to thought, and giving up the identification of thinking with mental representation. We take it that a cognitivist account of linguistic behaviour should take the form of an attempt to model and identify the states of and processes in (and around) the brain that are causally necessary conditions of linguistic behaviour. To identify such states and processes would be to provide an explanation of behaviour by appeal to obscure inner states and processes; so it would conform to the cognitivist brief. But to discover the causally necessary enabling conditions of linguistic behaviour is not to discover what thoughts are and what thinking is – for thoughts are expressed in behaviour, but a causal condition is not expressed in what it causally enables (see Trigg & Kalish 2010 [submitted]).

It is all too easy to confuse thinking, which speaking and writing is often expressive of, with the brain events and processes that are causally necessary conditions for intelligent linguistic behaviour. A Cartesian conception of the relation between thought and its linguistic expression makes such confusion almost inevitable, for in characterising thought as an inner accompaniment to observable linguistic behaviour it lumps it in with the various neural events and processes which are causally necessary for linguistic behaviour. But, Wittgenstein reminds us, intelligent speech does not consist in a series of observable movements of the face and throat on the one hand, and a series of hidden mental events and processes on the other. Rather, thought is in intelligent speech in roughly the way that distress is in an anguished cry and amusement is in a spontaneous peel of laughter. So, contrary to the apparently innocent Cartesian intuition, a person’s thoughts are typically not hidden behind their words, but precisely revealed by them. If we remember this we will not so easily seek to identify thinking with the neural conditions of speech, for it is obvious that these conditions are not revealed, as a person’s thoughts typically are, by the things they say. Of course these neural conditions are not hidden behind the linguistic behaviour they causally enable in the dramatic way Cartesian thoughts are supposed to be hidden behind some of the noises people make, for they count as hidden only because they are in the thinker’s skull, not because they are in the thinker’s mind.

Cognitive science, on our view, has the job of framing and testing hypotheses about the causally necessary enabling conditions of familiar psychological phenomena like thinking, imagining, remembering and willing. As long as these familiar phenomena are not conceived as private, inner accompaniments to observable behaviour and then identified with unfamiliar operations on mental representations, cognitive science can make a substantial contribution to our understanding of them. There is even a sense in which it might be appropriate to characterise neurological states, structures, events and processes which causally enable familiar phenomena like assertoric thinking (judgement) as mental representations, since without them it would be impossible for people to represent the world in the way that they do.

Initial Clarifications – Mental representation and the language of thought

If one holds that cognition is mental representation, and that mental representation consists in combinatorial operations performed on mental representations, then one seems compelled to accept that cognition is a quasi-linguistic operation. The basic idea here is that to combine representations to form complex ones is to perform syntactic operations analogous to those involved in combining words and sentences to form more complex sentences. (Fodor, 1975, 2008) In both cases, it may be supposed, the representational properties of complex representations will depend in a systematic way on the representational

---

2 E.g., Hauser, Chomsky & Fitch (2002) say, “In the varieties of modern linguistics that concern us here, the term “language” is used quite differently to refer to an internal component of the mind/brain (sometimes called “internal language” or “I-language”). We assume that this is the primary object of interest for the study of the evolution and function of the language faculty,” p1569
properties of the simpler ones. Thus a finite stock of elemental representations together with a finite number of syntactic rules for the combination of these, will yield an unlimited number of possible complex representations. This picture, which is surely very widely accepted (see, e.g., Schneider, 2010), seems to require what Fodor famously called a ‘language of thought’. The thesis that mental representation, like spoken and written representation, is linguistic, still plays an important, though poorly defined role in cognitive science. It is not hard to see why it is so popular. For one thing it seems to provide a way of explaining the unlimited number of different intelligent things intelligent creatures can do in reference to an equally unlimited number of different representational states their minds or brains can be in. For another it seems to support a computational theory of mind or cognition, since the relevant combinatorial or syntactic operations seem well suited to be conceived as computational operations of a kind that might be run by something like an organic equivalent to a computer.

Whilst the hypothesis that there is a language of mental representation consisting, not of words and sentences but of mental representations is a natural way to develop the concept of mental representation, it is no more than that. It could be discarded and the notion of mental representation retained. So it is important to appreciate that the argument we present is directed at the very idea that the representation of different intelligent things intelligent creatures can do in reference to an equally unlimited number of different representational states their minds or brains can be in. For another it seems to support a computational theory of mind or cognition, since the relevant combinatorial or syntactic operations seem well suited to be conceived as computational operations of a kind that might be run by something like an organic equivalent to a computer.

The Mental Representation Argument

On a picture of the relation between thought and language that deserves to be called the classical picture the purpose of speech and writing is to make thought, which is essentially private and psychological, public and perceptible. The basic problem with this picture, however its details are worked out, is this: if, when Paul says ‘Ducks run.’ there are two types of operation he is performing (one mental, inner and private, one physical, outer and public), and two kinds of item he is operating on (one mental, inner and private, one physical, outer and public), we will have to explain what the relation is between these operations and these items. That is, we will have to explain how, on the one hand, the combining of mental representations stands to the combining of words; and on the other, how the combined mental representations stand to the combined words. It should be clear, at least with a little reflection, that the prospects of providing such an explanation look dim. If we conceive mental representations as components of a non-verbal language we will have to explain how non-verbal thoughts are to be translated into verbal utterances; if we conceive mental representations as mental images or pictures we will have to explain what it is to translate images into words. But such explanations are hopeless for at least three sorts of reason. There is no such thing as translating what is said in a mental language into something said in a public one, so no such thing as doing it correctly or incorrectly (try to compare your translation of a mental representation with the mental representation it is a translation of so as to check if you’ve translated it correctly). Neither is there any such thing as translating – as opposed to describing – an image. Finally, since any image can be described in indefinitely many equally faithful ways, images do not determine what does and doesn’t count as faithfully describing them. So being able to say what one is thinking cannot be a matter of being able to describe a certain mental image correctly; since any description of a particular image could count as a correct description of it, talk of the (or even of a) correct description of an image is empty.3

If these weighty considerations do not convince thus baldly presented, we can turn to Wittgenstein’s celebrated (misunderstood and neglected) private language argument to drive them home. This succeeds in showing, quite categorically, that when, e.g., Paul says ‘Ducks run.’ he does not perform two types of operation, one on mental representations and one on verbal representations, but only one. It shows this by showing that operating on representations must be a normative or rule-governed affair (the type of thing that can be done incorrectly or correctly) and that operations performed on items, events or states available only to the operator could not be a normative or rule-governed affair. According to our argument if one assumes that thinking is a combinatorial operation on representations one has to deny that thinking is mental representing: that is interesting, for it is precisely that assumption that has led so many cognitive scientists to conclude that it must be mental representing. What follows then, is an argument to the conclusion that thinking cannot be an operation on mental representations which exploits Wittgenstein’s argument that language cannot be private:

To think, in the sense under discussion, is to think that something or other is so. To do that is to have a thought or to make a judgement, such that the thought one has or the judgement one makes will be the thought or judgement that things are thus and so. If things are that way the thought one has will be true, if not, then false.

For the purposes of argument let us assume that to think (or to judge, doubt or suppose) that something is so is to perform a certain kind of operation on a certain kind of representation. This operation must be productive of further representations that have the characteristic of being evaluable for truth; so it must be an operation on representations (that may or may not be of a truth-evaluable type) that yields representations (that are of a truth-evaluable type). So, for example, it could be an operation on words or sentences that produces sentences. Such operations must be combinatorial. What other than a combinatorial operation could produce the type of representation that is

---

evaluate for truth out of representations that are, very often, not evaluate for truth; and what other than a combinatorial operation could generate indefinitely many truth- evaluate representations out of a few non-truth evaluate ones?

Not every possible combination of representations will yield representations capable of truth. For example ‘Green was a depressing silently were or jam’ is a combination of representations, but it is nonsense, so the question of its truth cannot arise. There are many different ways of combining representations in such a way as to produce nonsensical representations not capable of truth. To combine representations in at least some of these ways that produce nonsense is to combine representations incorrectly. What can be done correctly or incorrectly must be a rule- governed activity. That is to say that for representations to be combined in the relevant way is not just for a series of orderly events to occur, but for a rule-governed activity to be engaged in. When a given process does not unfold as it usually does we can say that an irregularity has occurred in it but not that a mistake has been made in carrying it out. But many nonsensical combinations of representations are not just irregular; they are wrong. That entails that they run counter to rules of representation combination that do not merely capture actual regularities exhibited by representation combining activities but prescribe how those activities should be carried out.

Now if a given representation is knowable as the representation it is only by whatever it is that performs combinatorial operations on it, the rules determining how it may be combined with other representations must be private rules. There can’t be public rules that determine how private items, events or states should be manipulated, because it would be impossible to assess putative observations of such rules for correctness. If the rule is, “Perform operation p when y-type items appear, or x-type events occur, or r-type states are actualised,” and if y-type items, x-type events and r-type states are knowable only to the performer of p, then performances of p cannot be publically checked or assessed for correctness. So they can be checked for correctness only privately.

Now it seems that mental representations must precisely be representations knowable as what they are only by whichever thinker or representer is operating on them. To say that a given representation is a mental representation is to say that the dealings thinkers have with it are not perceptual. For a mental representation to be available to a thinker need not require that a certain publically available mark be seen or sound heard, but only that thinking be going on. Thinking is conceived here precisely in contrast to observable behaviour, as an inner or psychologival operation. Such an operation must be an operation on items, structures or states knowable as the representations they are only to the relevant operator.

If that is right, and if there can’t be public rules for the combination of private representations, then the rules that determine what it is to combine private mental representations correctly must be private rules.

Having reached this result, it only remains to be established that there can be no private rules for the combination of representations (or anything else), and it will have been shown that thinking cannot be a combinatorial operation on mental representations. Wittgenstein showed how to establish exactly that.

There can be no private rules governing the operations performed on mental representations because there can be no difference between its seeming to a thinker or representer on a given occasion that they are following a private rule and their really following a private rule on that occasion. Say Sam’s putative rule, p, is – “Whenever an item relevantly similar to this one (pointing inwardly to a relevant sample) comes before my mind, I will perform operation r on it (or whenever I am in this sort of state—pointing inwardly to an appropriate state—perform operation r on it)”. In these sorts of case there could be no difference between its seeming to Sam at t that the relevant item was before her mind, or that she was in the relevant state, and that item really being before her mind, or her really being in that state. In that case, Sam cannot have invented a rule p governing his performance of operation r, because any future performance that seems to Sam to be a performance of r in accordance with p, will thereby be a performance of r in accordance with p. In situations that do not allow for a distinction between what seems justified and what is, talk of being justified or unjustified is out of place.

To conclude, if thinking is a combinatorial operation on representations it must be a rule-governed combinatorial operation on representations; but now, since private rules for the combination of representations are impossible, and rules for the combination of mental representations would have to be private, thinking cannot be a combinatorial operation on mental representations. This argument shows that there is a fundamental conflict between the idea that thinking is an activity subject to normative constraint, and the idea that thinking is a private psychological affair: if thinking can be done incorrectly it cannot consist in manipulations of private mental representations.

Here is a concise formulation of the argument just given.

1. To think – in the relevant sense – is not just to think of something but to think that something is the case.
2. To think that something is the case is to combine representations that may or may not be the sort of representations capable of truth, so as to produce representations that are the sort of representations capable of truth.
3. It is possible to combine representations so as to produce representations that are not capable of truth as well as those that are.
4. To combine representations in such a way as to produce representations not capable of truth is to combine representations incorrectly.

4 See Wittgenstein (1958) e.g. section 258.
5. If something, \( p \), can be done incorrectly there must be rules that determine what counts as doing \( p \) correctly; \( p \) must be a rule-governed activity.
6. By 2, 3, 4 and 5, thinking must be a rule-governed activity. (Not just a process exhibiting regularities).
7. Rules for the combination of representations knowable as what they are by only one representor would have to be private rules.
8. Representations that are mental must be knowable as what they are only by one representor.
9. So by 7 and 8 – rules for the combination of mental representations must be private rules.
10. There can be no private rules.
11. The combination of mental representations – by 7, 8, 9 and 10 – cannot be a rule-governed activity.
12. Thinking – by 6 - must be a rule-governed activity.
13. Thinking – by 11 and 12 – cannot be a combinatorial operation on mental representations.

An Objection

Without further ado, let us consider an objection to this argument. It concerns premise eight - the claim that mental representations, qua mental, must be private, that is, must be knowable as the representations they are only by whatever it is that represents by operating on them. Many cognitive scientists and philosophers of mind might eagerly reject this premise on the ground that it depends upon an unacceptably Cartesian notion of what mental representations are. It will be said that if we reject this out-dated Cartesianism, and think of the mental representations in question as states of, or items in, brains, (or as realised in such items or states), we can deny that they are private, and so allow room for the idea that there could be public rules for their combination.

The problem with this objection is that it is inconsistent with the claim that thinking is mental representing, so can’t be used to defend it. The claim that thinking is mental representing depends on a certain way of conceiving the distinction between representing done in thought and representing done in (public) language. On this conception representing done in thought (mental representing) is what makes representing done in public language what it is; it is because Sam’s assertions do, but his sneezes do not, depend somehow on his thoughts, that his assertions count as meaningful utterances rather than mere noises. It turns out, we will now argue, that this conception allows for the possibility that a representor may be wrong about the assertion they are making, but it excludes the possibility that a representor may be wrong about the judgement they are making (the thought they are having). We argue that rejection of premise 8 is incompatible with the Cartesian view that a thinker cannot be wrong about which judgement they are making, and that this Cartesian view is an essential feature of the position rejection of this premise is meant to defend. Thus, whilst the anti-Cartesian feel of the objection now under discussion may be congenial in itself, it is quite inconsistent with the conception of the relation between thought, mental representation and language that it is meant to defend.

Speakers are sometimes wrong about what they are saying. This is not to say that speakers are sometimes insincere; it is to say that speakers can think they are saying one thing when they are really saying another. How shall we explain the possibility of this sort of mistake?

If we hold that to think is mentally to represent, and that thoughts are mental representations we will have to hold that to make a sincere assertion is to translate or otherwise convert a mental representation into a non-mental one. This commits defenders of this conception of thinking to the view that for a speaker to be wrong about what they are saying is for a speaker to be wrong about the relation between what they are saying and what they are thinking. Mistakes of that kind are mistakes made in translating or converting mental into non-mental representations. Now, on this conception of how it is possible for a speaker to be wrong about what they are saying, no conceptual room is left for the possibility that they could be wrong about what they are thinking. This is because, on this view, whilst a speaker has to convert or translate their thoughts into sentences in order to make an assertion, they do not have to translate or convert their thoughts into anything in order to have them. If one can be wrong about what one is saying because one can translate or convert one’s thoughts into words incorrectly, and one does not have to translate or convert one’s thoughts into anything in order to have them, one cannot be wrong about what thought one is having at a given time.

To see this, consider the following argument. If we explain what it is for someone to make a mistake about what they are saying by appeal to the idea that they can be wrong about the relation between what they are saying and what they are thinking, we will have to deny that they can be wrong about what they are thinking, on pain of generating an infinite regress. For if it were possible for a thinker, Sam, to be wrong at \( t \) about what she is thinking at \( t \), that possibility would require explanation. Any such explanation would have to appeal to a difference between what Sam is thinking at \( t \) and what she thinks she is thinking at \( t \) – we will have to say that whilst she thinks she is thinking one thing she is really thinking another. But as soon as we say that, we will also have to allow the possibility that there can be a further difference between what Sam thinks she is thinking at \( t \) and what she thinks she thinks she is thinking at \( t \), and so on. This regress is by no means benign, for it requires a thinker to have an infinite number of appropriate thoughts at \( t \) if they are to know what they are thinking at \( t \).

As soon as we open an anti-Cartesian gap between what Sam is thinking at \( t \) and what she thinks she is thinking at \( t \), for her to know what she is thinking at \( t \) it will not be enough that she thinks it at \( t \). Suppose she thinks that ducks run at \( t \). If she is to know that she is thinking that ducks run at \( t \), she has to think that she is thinking that ducks run at \( t \), and if she is to know that, she has to think that she is thinking that she thinks that ducks run at \( t \), and so on.
So if we are to avoid this regress we have to embrace Cartesianism; we have to say, that is, that Henry’s knowledge of what he is thinking is both incorrigible and evident; incorrigible because if he thinks he is thinking that p he is thinking that p, and evident because if he thinks that p he thinks that he thinks that p. If we take thoughts to be mental representations and assertions to be translations of mental representations into perceptible signs, we commit ourselves to an explanation of how a speaker can be wrong about what they are saying which only Cartesianism will save from incoherence.

Now of course the relevant point is that this Cartesian account of the relation between a thinker and the thoughts they have is flatly incompatible with the proposal that mental representations are not private. To see this, it is important to appreciate first that to say that a given mental representation is publically available is not just to say that it is identical with certain brain states that are publically available. It is conceivable that Sam should be acquainted with a certain brain state, p, which is in fact identical to a certain mental representation r, but know neither that p is a mental representation nor that p is mental representation r. (If I know the butcher, and the butcher is the president, then I know the president, but I may not know that the butcher is the president). So what is required is that George and Harry can come to know that Grace is thinking that p by becoming acquainted with a certain state of Grace’s brain.

Now that possibility is rather dramatically incompatible with the Cartesian conception of thinking to which we have just shown our opponent to be committed. It makes Grace’s way of finding out what thought she is currently having into just one of many ways of finding that out. So a situation will be conceivable in which Grace tries to find out what she is thinking using her introspective method, George and Henry do the same by observing her brain, and Grace fails whilst George and Henry succeed. If that is thinkable, then not only could Grace be wrong, and George right, about what Grace is thinking at any time, but Grace could be wrong and George right about what she is thinking at all times!

The problem remember is not just that these possibilities are absurd in themselves – although the idea that something could count as a thinker whilst always being wrong about what it thought is pretty unsatisfying all on its own – but that they are incompatible with the Cartesian conception of the relation between thought and language to which our opponent is committed.

So the thesis that to think is to operate on mental representations cannot be defended by rejection of premise 8. If mental representations are constitutive of thoughts they must be private, and if mental representations are private combining them cannot be a rule-governed operation.

Conclusion

We take it that the mental representation argument shows that thinking cannot be mental representation, that is, that it cannot be a combinatorial operation on mental representations. While there are objections to this argument that we have not explicitly considered here, they will have to turn either on the denial that thinking can be done incorrectly or on the claim that there can be private rules, and these responses seem to head off in unpromising directions. The only plausible option open to the representationalist is to conceive thinking as a combinatorial operation on the representations constitutive of a natural language like English (Malcolm.1973). We have said nothing either for or against that position here – though it is perhaps worth noting that anyone attracted to it will have to hold that it is persons as we ordinarily conceive them and not minds or brains that think (since it is indubitably human beings and not minds or brains that know how to use the words of a natural language).

So thinking is not mental representing, and, for example, asserting is not converting or translating mental into non-mental representations. Does this result show that a cognitive scientific account of linguistic behaviour is impossible? Not at all. It shows that if we identify mental representing with thinking we cannot explain intelligent linguistic behaviour by appeal to mental representing. But if we think of mental representing not as identical to thinking but as a causally necessary condition on it (Trigg & Kalish, 2010), then the idea that mental representation underlies intelligent linguistic behaviour is in good shape. It is, of course, profoundly plausible that if certain very complex events did not take place in a person’s brain at t they would not be able to think or speak at t. Nothing in the argument just presented is incompatible with this idea, and nothing suggests that the difficult business of finding out about these events is not a scientific undertaking of the greatest interest.

References


