Margolis, E., & Laurence, S. (1999). Where the Regress Argument Still Goes Wrong: Reply to Knowles. *Analysis*, 59.4, pp. 321-327.

The final version is available at Oxford University Press / JSTOR:
https://www.jstor.org/stable/3328603

Where the Regress Argument Still Goes Wrong: Reply to Knowles

Eric Margolis and Stephen Laurence

Abstract. Many philosophers reject the Language of Thought Hypothesis (LOT) on the grounds that is leads to an explanatory regress problem. According to this line of argument, LOT is invoked to explain certain features of natural language, but the language of thought has the very same features and consequently no explanatory progress has been made. In an earlier paper ("Regress Arguments against the Language of Thought", *Analysis* 57.1), we argued that this regress argument doesn't work and that even proponents of LOT have given the regress argument too much credit. In this paper, we extend our critique of the regress argument against LOT by responding to J. Knowles's commentary of our earlier paper (in "The Language of Thought and Natural Language Understanding", *Analysis* 58.4).

1. Introduction

The Language of Thought Hypothesis (LOT) is at the centre of a number of the most fundamental debates about the mind. Yet many philosophers want to reject LOT out of hand on the grounds that it is essentially a recidivistic doctrine, one that has long since been refuted. According to these philosophers, LOT is subject to a devastating regress argument. There are several versions of the argument, but the basic idea is as follows. (1) Natural language has some important feature, X.[1] (2) Defenders of LOT appeal to an internal system of representation in order to explain this feature of natural language. (3) Yet the hypothesized language of thought also has X. (4) This raises the following dilemma: If we offer an analogous explanation of the language of thought's having X, we are off on a regress. If we offer some other explanation, then the alternative explanation should have been given for natural language in the first place, avoiding the detour through the language of thought.

In Laurence and Margolis (1997), we argued that even the most ardent supporters of LOT have been too generous in their responses to the Regress Argument; when seen in the proper light, the Regress Argument itself does nothing at all to impugn LOT. The crucial point is that the Regress Argument gains its rhetorical force from the tacit assumption that the only point to positing a language of thought is to address the explanandum cited in the regress. For suppose that there are independent motivations for positing a language of thought – that is, motivations apart from explaining how natural language has feature X. In that case, the LOT theorist is free to appeal to

[1] Different candidates for feature X include that language is learned, that it is understood, and that it is meaningful.

the language of thought in accounting for natural language possessing feature X, *but she is not required to*.

Presumably, she should appeal to the language of thought if doing so has theoretical advantages over explaining feature X of natural language directly. And it can be argued that explaining feature X by reference to the language of thought *does* have many advantages, even though the language of thought itself has feature X. LOT theorists have jumped on this point, effectively challenging the second horn of the dilemma. However, this response grants far too much to the regress argument. For the LOT theorist is equally free to explain feature X without appealing to the language of thought, since there is independent reason for adopting LOT. So, for example, it might turn out that the meaningfulness of natural language and the meaningfulness of the language of thought are each explained directly and independently of one another (more on this below). Contrary to what proponents of the regress argument claim, LOT is not 'destroyed' in this situation.

In short, for the Regress Argument to have any force at all, it would have to be accompanied by further arguments that undermine all of the independent motivations for LOT. But in that case, the Regress Argument would be an inconsequential part of a much more general argument to the effect that LOT is unmotivated. By itself the Regress Argument poses no threat to LOT.

2. Knowles's Critique of Our Solution

Jonathan Knowles (1998) disagrees with our assessment of the Regress Argument.[2] Unfortunately, his response is based on a number of misunderstandings of our position and the dialectical situation surrounding LOT. Knowles identifies two central theses in our paper, both of which he aims to refute. These are (1) that the Regress Argument has 'been answered incorrectly or inadequately by supporters of LOT' (263), and (2) that LOT is not undermined anyway since 'the main sources of support for LOT are (or might be) independent of its providing an account of how we understand natural language' (263). Knowles's discussion of each of these theses is problematic. We'll take them up them in reverse order.

Though we never claim (2), it is probably true that the main sources of evidence for LOT are independent of considerations pertaining to natural language. In any event, Knowles's claim that natural language is 'necessarily a central part of the support for LOT' (264) is certainly false. This is clear from the fact that non-linguistic creatures also have cognitive abilities that warrant the postulation of a language of thought. We have in mind non-human animals, prelinguistic humans, aphasiacs, and other individuals who, for whatever reason, lack a natural language. Individuals in these various categories are often capable of engaging in sophisticated problem solving, various types of decision making, and other tasks that would seem to require an internal system of representation.

[2] All page references are to Knowles (1998) unless explicitly noted.

Consider, for example, the impressive capacity of certain species to successfully store and retrieve food caches. It's been estimated that the Clark's nutcracker needs to retrieve approximately a thousand caches in order to survive a typical winter. Studies show that this ability doesn't depend upon olfactory information or rigid searching routines. Rather, Clark's nutcrackers are in possession of a rich, flexible cognitive map. In other words, they invoke a system of representation that records the spatial layout of their environment together with pertinent information that can be used to mark the specific locations of caches (see, e.g., Griffin 1992). Cognitive maps are hardly unique to birds. Even creatures with as little neural structure as bees and ants have impressive abilities that turn on this kind of representation. In a review of some of the feats of which insects are capable – sun compass course holding, dead reckoning, determining distance from parallax, etc. – the psychologist Charles Gallistel concludes that insects must rely on a symbol-processing system (Gallistel 1998). Of course, it's an empirical question whether such a system would have to have a compositional syntax and semantics, but in the same way it's an empirical issue whether such a system is needed in the case of language.[3]

We don't deny that LOT theorists often do cite natural language as a motivation for LOT, or that they should. Language involves some of our most impressive and best understood cognitive abilities. Moreover, language introduces a clear case of a productive and systematic capacity, and productivity and systematicity are at the heart of the empirical case for LOT. At the same time, however, there simply is no reason to hold, as Knowles does, that the LOT is *necessarily* tied to natural language understanding.[4]

Regarding (1), Knowles's interpretation of our position is again distorted. According to Knowles we claim 'by implication' that 'in the absence of other support for LOT, those who put forward a naturalistic theory of intentionality for LOT as a way of explaining natural language meaning and understanding gratuitously assume that such theories must apply to the internal medium rather than to natural language' (266). But this is not our position at all. Rather, our view is that theorists who choose to respond to the Regress Argument by taking on the second horn of the dilemma owe some account of why it's not gratuitous to invoke the language of thought in their explanations of natural language. And far from claiming that no such account would be forthcoming, we've even explained what might in fact motivate the detour through LOT, adding that we have 'considerable sympathy' with this strategy (Laurence & Margolis 1997, 63). Our point is simply that there is a *stronger* response to the regress argument above and beyond this one. For even if this sort of account isn't available, the regress argument would still fail to pose a serious threat to LOT. LOT isn't invoked merely to explain feature X of natural language.

---

[3] Moreover, for purposes of evaluating the Regress Argument, we needn't be concerned about these features of LOT, since the argument is supposed to apply to any theory that invokes an internal system of representation whether it has language-like structure or not.

[4] The name 'Language of Thought Hypothesis' may be a bit misleading, since it may appear to mention language (i.e., natural language). Other theorists have been confused by this point. E.g., Patricia Churchland in a critical discussion of LOT refers to the 'infralinguistic catastrophe', that is, the problem that 'intelligent behaviour is displayed by organisms who have no overt linguistic capacity' (Churchland 1986, 388). But LOT theorists have always been clear that LOT isn't a thesis about natural language. For instance, in Jerry Fodor's seminal discussion of LOT (his 1975), he introduces the hypothesis without so much as mentioning natural language (see chapter 1). Instead he motivates the language of thought by reference to three central explananda in psychology which are independent of natural language processing – practical reasoning, concept acquisition, and perception.

3. A Better Solution?

Let's now turn to Knowles's own response to the Regress Argument, which centres around two claims. First, he argues that the regress on learning is best met by holding that the language of thought is innate (following Fodor 1975). Second, he claims that an adequate general response to the Regress Argument can be made by providing an account of meaning for LOT, given that there are reasons why this account can't be applied directly to natural language. We briefly comment on these in turn.

Knowles's response to the regress on learning is rather surprising, as Fodor's notoriously strong nativism (which Knowles seems happy to endorse) has been nearly universally rejected. Nonetheless, Knowles claims that innateness isn't 'merely an optional extra of LOT, for it *is* problematic how we could learn any language without some prior understanding. It seems at least correct to say that to the extent that the language of thought can be seen as innate, LOT is strengthened...' (267; italics in original). Of course, it is certainly true that if the language of thought is innate, then there must be a language of thought. And it is also true that the recourse to innateness does address some of the problems raised by the Regress Argument. But the innateness thesis is no easy pill to swallow. Fodor's nativism specifically requires the innateness of almost every lexical concept that any human has ever had or will have (including, e.g., the concepts of triangle, trumpet, bachelor, carburettor, electron, and cigar.) This is a wildly implausible doctrine, to say the very least. Some have even gone so far as to claim that Fodor's innateness thesis itself constitutes a *reductio ad absurdum* of LOT (see Churchland 1986).[5] So Knowles is hardly doing LOT a favour by burdening it with this extra commitment.

What about the second half of Knowles's response to the Regress Argument? Like many who are sympathetic to LOT, Knowles claims that naturalistic theories of intentionality are suited to apply directly to the language of thought, not to natural language. In fact, Knowles goes one step further. He claims that 'arguments for LOT *presuppose* the rejection of the idea that linguistic meaning is based on relationships between words in a natural language and/or between these words and their extensions, on the grounds that such relationships will not be naturalistically explicable' (268; italics in original). That is, according to Knowles, naturalism is an intrinsic feature of LOT, and naturalism requires that natural language meaning be explicated in terms of meaning in the language of thought.

This is a strong claim and one that is simply unwarranted. Whether the language of thought appears to be the better candidate for direct naturalization depends on the theory of content one adopts. For instance, in our earlier article we cited Fodor's remark that mental representations aren't under

[5] Though Knowles isn't explicit about exactly what's innate on his view, it's a plausible reading of the text that he does endorse Fodor's radical nativism. Strictly speaking, however, Knowles could adopt a weaker doctrine by maintaining that lexical concepts are structured and that only a relatively small stock of primitive concepts are innate – the rest being assembled from their constituents. But theories in this tradition face a number of daunting criticisms (see Laurence & Margolis 1999).

voluntary control in the way that natural language symbols are. That makes them far better targets for nomic co-variation theories of content (such as Fodor 1990), since mental representations are more likely than words to be tokened as a reliable effect of causal interaction with things in their extension. (When you see a chair, maybe you can't help but think 'chair', even if you don't say it.) So if one is independently committed to a nomic co-variation theory of content, then a reduction from natural language to the language of thought may very well be in order. But arguments for the language of thought don't actually presuppose any particular theory of content. That's why even among supporters of LOT the nature of mental content remains a hotly debated topic. As a result, a reduction from natural language to the language of thought is hardly mandatory.

Take, for example, Ruth Millikan's theory of content, which relies quite heavily on an evolution-based notion of teleology (Millikan 1984). One of the most interesting features of this theory is precisely that it doesn't involve a reduction of linguistic content to thought content. As she develops the theory, the functions assigned to beliefs receive one treatment, desires another, and importantly, different types of linguistic devices (indicative sentences, imperatives, etc.) each receive separate treatment as well. Language does not reduce to thought; rather, language and thought are given parallel treatments. To be sure, Millikan's discussion isn't framed in terms of LOT, but that's irrelevant. There is no reason why a LOT theorist couldn't adopt the essence of her theory, giving independent teleological accounts of the content of language of thought expressions and the content of natural language expressions. This isn't to say that we are endorsing a model of this sort, or any particular theory of content. For present purposes, all that matters is that it is a real theoretical possibility that the content of natural language isn't inherited from the content of the language of thought. This shows that Knowles is simply wrong to suppose that LOT requires that natural language meaning reduces to meaning in a language of thought.

4. Conclusion

Knowles's solution to the Regress Argument is flawed in a number of ways. He wrongly claims that LOT is intrinsically connected to natural language use, he endorses Fodor's extremely implausible radical concept nativism, and he mistakenly claims that LOT presupposes that the intentionality of natural language must derive from the intentionality of the language of thought. What's more, his criticism of our alternative response misconstrues our position. The result is that much of his discussion is devoted to defending a claim that we have never disputed, viz., that LOT theorists can challenge the second horn of the dilemma in the Regress Argument. We agree that they can, and we've explicitly expressed sympathy with this line of argument. Yet it still doesn't get to the heart of the matter. For even if the standard response to the Regress Argument were not available, the Regress Argument *still* wouldn't have any force against LOT since there are independent motivations for LOT. In order to undermine LOT, the proponent of the Regress Argument would have to provide additional arguments against all of these sources of evidence. Seen in this light, the regress argument constitutes at best a more or less insignificant component of a much more general – and far harder to mount – argument to the effect that there isn't any

positive evidence for LOT. On its own, the Regress Argument is inconsequential to the evaluation of LOT.[6]

References

Churchland, P. 1986. *Neurophilosophy*. Cambridge, Mass: The MIT Press.

Fodor, J. 1975. *The Language of Thought*. Cambridge, Mass: The MIT Press.

Fodor, J. 1990. A theory of content, II: the theory. In *A Theory of Content and Other Essays*. Cambridge, Mass: The MIT Press.

Gallistel, C.R. 1998. Symbolic Processes in the Brain: The Case of Insect Navigation. In Scarborough and Sternberg (eds.) *An Invitation to Cognitive Science*, Volume 4: Methods, Models, and Conceptual Issues. Cambridge, Mass: The MIT Press.

Griffin, 1992. *Animal Minds*. Chicago: University of Chicago Press.

Knowles, J. 1998. The language of thought and natural language understanding. *Analysis* 58.4: 264–72.

Laurence, S., and E. Margolis. 1997. Regress arguments against the language of thought. *Analysis* 57: 60–66.

Laurence, S., and E. Margolis. 1999. Concepts and Cognitive Science. In Margolis and Laurence (eds.) *Concepts: Core Readings*. Cambridge, Mass: The MIT Press.

Millikan, R. 1984. *Language, Thought, and Other Biological Categories*. Cambridge, Mass: The MIT Press.

[6] This paper was fully collaborative; the order of the authors' names is arbitrary.