

*The Compositionality Papers*, by Jerry A. Fodor and Ernie Lepore. Oxford: Clarendon Press, 2002. Pp. Viii + 212. \$19.95.

Zoltán Gendler Szabó  
Sage School of Philosophy  
Cornell University  
[zoltan.szabo@cornell.edu](mailto:zoltan.szabo@cornell.edu)

Forthcoming: *Mind*

Compositionality matters, or so Fodor and Lepore repeatedly tell us on the pages of this collection of essays. It matters because it is undeniable that human thought and language are compositional and because this fact imposes severe constraints on theories of mental and linguistic representation. As they put it: ‘It’s our belief that, of the various familiar candidates for concepthood, only very few – perhaps only one – can meet this compositionality condition; it’s the bull in everybody’s china shop. If that’s right, then compositionality tells us what concepts and word meanings are.’ (p. 4)

This is surprising, if true. Not only is compositionality widely accepted among linguists, psychologists and philosophers who disagree about meaning, it actually strikes many as a rather modest, perhaps trivial requirement. There is a sense in which just about any meaning-assignment can be compositionally imitated. Perhaps the best known formal result illustrating this is the following: given a set  $\underline{S}$  of strings generated from an arbitrary alphabet via concatenation and given a meaning function  $\underline{m}$  which assigns to the members of  $\underline{S}$  members of an arbitrary set  $\underline{M}$ , we can construct a new function  $\underline{\mu}$  such that for all  $s, t \in \underline{S}$   $\underline{\mu}(s.t) = \underline{\mu}(s)(\underline{\mu}(t))$  and  $\underline{\mu}(s)(s) = m(s)$ . (Zadrozny, W. (1994) ‘From Compositional to Systematic Semantics.’ *Linguistics and Philosophy*, 17: 329 - 342.) To be sure, this result has not kept many semanticists awake at night. But that’s because the values of  $\underline{\mu}$  in Zadrozny’s proof are fanciful functions definable only within the theory of

non-wellfounded sets and nobody believes that those could really be the meanings of ordinary English expressions. In other words, we need independent constraints on what meanings could be if we want compositionality to have a bite. But then how could compositionality *by itself* tell us anything important about the nature of meaning?

One source of this divergence is that Fodor and Lepore have a stronger notion of compositionality in mind than Zdrozny does. Compositionality, as typically understood by semanticists, is the claim that the meaning of a complex expression is a *function* of the meanings of its constituents and the syntactic way these constituents are combined. By contrast, Fodor and Lepore interpret compositionality as the principle that the meaning of a complex expression is *determined by* (pp. 29, 59, 112, 143, 176, 179), *inherited from* (pp. 1, 4, 14), or *constructed from* (p. 43) the meanings of its constituents and the syntactic way these constituents are combined. Functions are cheap. Suppose we systematically assign natural numbers to all (disambiguated) English sentences, different numbers to different sentences. Then the meaning of any English sentence is a function of its associated number. Nonetheless, given that sentences mean what they do independently of our assigning this or that natural number to them, the meaning of an English sentence is not determined by, inherited from or constructed from its associated number.

Keeping in mind that Fodor and Lepore (quite sensibly) take compositionality to be a stronger claim than many formal semanticists do helps in appreciating some of their arguments. Take Fodor's early argument against prototype theory, reiterated and expanded in the second essay of the collection. We know what the prototype of a pet is (say, a dog) and what the prototype of a fish is (say, a carp). But it does not seem that

from these (together with what we know about the syntax of adjectival modification) we could figure out what the prototype of a pet fish is (say, a goldfish). Surely, this has considerable intuitive appeal, but not because anything has been said to cast doubt on the existence of a function from dogs, carps and the syntax of adjectival modification to goldfish. The relevant intuition is rather that the prototype of a pet fish could be different even if the prototypes of pet and fish (as well as the syntax of English) were to remain the same, and that consequently, the former cannot determine the latter.

Whether focusing on their strong conception of compositionality can save their argument against inferential role semantics is less clear. Here is how they argue that total inferential roles are not compositional: “Suppose, for example, that you happen to think that brown cows are dangerous; then it is part of the inferential role of ‘brown cow’ in your dialect that it does (or can) figure in inferences like ‘brown cow’ → ‘dangerous.’ But, first blush anyhow, this fact about the inferential role of ‘brown cow’ doesn’t seem to derive from corresponding facts about the inferential roles of its constituents.” (16) I don’t see why not. If it is part of the inferential role of ‘brown cow’ that ‘brown cow’ can feature in inferences like ‘brown cow’ → ‘dangerous’, why would it not be part of the inferential role of ‘brown’ that ‘brown’ can feature in the very same inferences? (In fact, why would we call any inferential role associated with ‘brown’ that does not reflect my belief that brown cows are dangerous *total*?) Fodor and LePore may, of course, insist that the inferential roles of complexes are not inherited or constituted by the inferential roles of their constituents, but it is hard to avoid the feeling that this would require some independent justification. I myself think that although there are excellent arguments against the identification of meanings and (total) inferential roles, many of which were

presented by Fodor and LePore in their 1992 book *Holism, A Shopper's Guide* (Oxford: Blackwell), the argument from compositionality is not one of them.

Be that as it may, the example of prototypes already illustrates that compositionality, if construed appropriately, eliminates certain conceptions of meaning. But does it eliminate conceptions of *lexical* meaning? According to Fodor and Lepore, 'it is hard to see, for example, how primitive meanings could be birds or chairs since, whatever complex meanings may be, it's hard to see how birds or chairs could be parts of them.' (p. 57) But is it really? Let's say that the meanings of complex expressions are interpreted logical forms, i.e. phrase structure trees with the meanings of the constituent lexical items assigned to their terminal nodes. I take it that in a not-too-farfetched sense the meanings of lexical items are then parts of the meanings of complex expressions in which they occur, and it would even be sensible to say that the meanings of complexes are determined, inherited from, or constructed from the relevant birds and chairs. So how could compositionality, however strongly one wishes to interpret it, serve as 'the sovereign test for theories of lexical meaning'? (p. 43)

There are a number of places where Fodor and Lepore are quite explicit about why they think it is. They claim that compositionality requires the *context-independence* of lexical meanings. That is, they think that 'the fact that 'brown dog', 'green dog', 'brown cat', and 'yellow cat' are all compositional in English is part and parcel of the facts that 'brown' means the same in the environment '...dog' that it does in the environment '...cat'; 'dog' means the same in the environment 'brown...' that it does in the environment 'green...'; and so on.' (p. 180) But this does not follow from compositionality. No matter how much the meanings of lexical items may vary with the

linguistic environment in which they occur (or, for that matter, with extra-linguistic factors), as long as all the context-dependence in complex expressions can be traced back to context-dependence in their lexical parts, compositionality remains unchallenged. The meanings of parts (in context) still determine the meanings of the wholes (in context).

Interestingly, at one point Fodor and Lepore are more careful about where they think the context-independence requirement is coming from: ‘It’s pretty widely agreed that an explanation of the fact that complex meanings are systematic requires assuming that lexical meanings are context-independent. The idea is this: compositionality says that the meanings of ‘John snores’ and of ‘John swims’ depend *inter alia*, on the meaning of ‘John’. And it’s because ‘John’ means the same in the context ‘...snores’ as it does in the context ‘...swims’ that, if you know what ‘John’ means in one context, you know thereby what it means in the other.’ (p. 58) That is, context-independence of meaning – undoubtedly a highly substantive constraint on lexical semantics – is not a consequence of compositionality, but a principle that along with compositionality plays a significant role in explaining systematicity.

Here is my conjecture. What really does the job of constraining what lexical meanings might be (although perhaps not as much as Fodor and Lepore think) is not compositionality but rather whatever best explains productivity and systematicity. (They do in fact say things along this line in the third essay of the collection.) Compositionality is one plank in such an explanation but there are many others. We already saw that postulating context-independence of lexical meaning is also important. In addition, we must assume the finiteness of the lexicon, the relative simplicity of syntax, and probably much more.

If it is the best explanation of productivity and systematicity that is supposed to underlie the proposed constraints on theories of meaning, many of the arguments in this book begin to sound rather plausible. (I concede that the more appropriate title – *The Best Explanation of Productivity and Systematicity Papers* – would have been less catchy). Take the argument against the identification of meanings with inferential roles I criticized above. The standard explanation for why knowing the meanings of ‘black dog’ and ‘white cat’ is sufficient for knowing the meanings of ‘black cat’ and ‘white dog’ is that understanding complex expressions requires the ability to decompose them into their lexical constituents and then recombine those primitive meanings into the meanings of other complex expressions. But if this is indeed something we can do, then it better be the case not only that the meanings of lexical items (together with the syntax) determine the meaning of any complex expression those lexical items compose, but also that the meaning of any complex expression determines the meanings of its lexical constituents. In other words, the standard explanation of systematicity requires not only compositionality, but also what Fodor and Lepore call *reverse compositionality*. (p. 59) And even though inferential roles are compositional, they violate reverse compositionality. The inferential role of ‘brown cow’ does not determine the inferential role of ‘brown’: you and I may be inclined to draw the same inferences essentially involving ‘brown cow’ even if you think brown dogs are dangerous and I don’t, and consequently we differ in our inferential roles of ‘brown’.

If you’re joining the review here at the last paragraph let me assure you that what is said above gives you reason to go out and buy this engaging, irreverent and

provocative book. Most of its essays are already in the focus of debate and they will surely remain there in years to come.