ABSTRACT. In the context of debates about what form a theory of meaning should take, it is sometimes claimed that one cannot understand an intersective modifier-head construction (e.g., ‘pet fish’) without understanding its lexical parts. Neo-Russellians like Fodor and Lepore contend that non-denotationalist theories of meaning, such as prototype theory and theory theory, cannot explain why this is so, because they cannot provide for the ‘reverse compositional’ character of meaning. I argue that reverse compositionality is a red herring in these debates. I begin by setting out some positive arguments for reverse compositionality and showing that they fail. Then I show that the principle of reverse compositionality has two big strikes against it. First, it is incompatible with all theories of meaning on the market, including the denotationalism favored by neo-Russellians. Second, it explains nothing that is not already explained by its venerable predecessor, the principle of (forward) compositionality.

It’s a familiar fact about speakers of English that anyone who knows the meaning of ‘pet’ and the meaning of ‘fish’, and has the syntactic wherewithal to combine these words, knows (at least implicitly) the meaning of ‘pet fish’. More generally, given any modifier-head construction – for example, any intersective noun phrase – one can understand any of its instances whose lexical parts one understands. This epistemological fact, a fact about speakers’ knowledge of meaning, is usually explained by appeal to the principle of compositionality (PC). The principle governs the nature of the rules knowledge of which underlies semantic competence (Larson and Segal, 1995). According to PC, the meaning of a modifier-head construction is derivable by these rules solely from the meaning of its lexical parts plus its syntax. No further information, either about language or about the world, is needed.¹ Compositionality in this sense is old hat, its theoretical importance beyond dispute. There’s a long-standing consensus in linguistics and philosophy, for example, that compositionality is needed to explain how natural languages can be learned, given that languages are infinite and brains are not (Davidson, 1984).
Lately it’s been argued that meaning in natural language is also compositional in a further sense, roughly converse to the traditional one (Fodor, 1998b, 2002; Fodor and Lepore, 2001). Following Fodor and Lepore, I’ll call this new property ‘reverse compositionality’. According to the principle of reverse compositionality (PRC), the meanings of a phrase’s lexical parts are derivable solely from the meaning of the phrase plus its syntax, and nothing more. Proponents of PRC claim that this new principle is needed to account for why anyone who knows the meaning and syntax of ‘pet fish’ also knows (at least implicitly) the meaning of ‘pet’ and the meaning of ‘fish’ – or, more generally, why anyone who understands a modifier-head construction also understands the words it contains, solely in virtue of understanding that one phrase. Like compositionality of the traditional variety, reverse compositionality is put forth as a structural property of language itself, posited to explain a feature of our knowledge of language. In each case, the linguistic property is a type of derivability, and its epistemic manifestation is a type of dependence in understanding.

Compositionality principles have loomed large in recent discussions of the theory of meaning. In particular, contemporary neo-Russellians – that is, advocates of a purely denotational (truth-conditional) approach to semantics – have pointed to them as knockdown evidence against rival views. Here is a representative passage:

Over the last few years, we have just about convinced ourselves that compositionality is the sovereign test for theories of lexical meaning. So hard is this test to pass, we think, that it filters out practically all of the theories of lexical meaning that are current in either philosophy or cognitive science. (Fodor and Lepore, 2001, p. 351)

The list of failures includes, *inter alia*, prototype theory, theory theory, and use theory. In fact, it includes all views except for two: the classical view of lexical meanings as definitions (discredited on other grounds) and the denotation-based account favored by the authors.

To see the test in action, consider how standard prototype semantics treats ‘pet fish’. According to such a semantics, the meaning of an expression is its associated prototype, that is, a list of features characteristic of items to which the expression correctly applies. The prototype of ‘pet’ includes features such as *furry,* *four-