On Mathematical Proofs of the Vacuity of Compositionality *

Dag Westerståhl
Stockholm University
dag.westerstahl@philosophy.su.se

February 23, 1998

1 Introduction

There is a widespread feeling among formal semanticists that, with suitable manipulations, any semantics can be made compositional. The following quote is representative: "...it [compositionality] being a methodological starting point, it is always possible to satisfy compositionality by simply adjusting the syntactic and/or semantic tools one uses, unless that is, the latter are constrained on independent grounds." (Groenendijk and Stokhof [4], p. 93). Sometimes this feeling is backed by a mathematical *proof*, purportedly showing compositionality to be vacuous, or without empirical content. A recent instance is W. Zadrozny's paper [9] in this journal, where he takes a fact about hypersets (non-wellfounded sets) to show that "...any semantics can be encoded as a compositional semantics, which means that, essentially, the standard definition of compositionality is formally vacuous." (p. 329). In this note I will examine this claim, as well as two earlier ones in the same vein.

2 A minimalist version of compositionality

Since Montague, compositionality has been expressed formally as the existence of a homomorphism from a syntactic to a semantic algebra, cf. Janssen [6], [7], and Hendriks [5] for elaborations. Here, the following account, which makes minimal algebraic assumptions, suffices:

On the syntactic side we have *expressions with structure*, represented as a *partial algebra* $\mathbf{A} = (A, (F_\gamma)_{\gamma\in\Gamma})$, where each F_γ is a partial operation on A of

*This work was partially supported by the Bank of Sweden Tercentenary Foundation in connection with the project *Meaning and Interpretation*. The author would like to thank Jon Barwise, Larry Moss and Peter Pagin for helpful comments.

a fixed (finite) arity. As for semantics, we have a *meaning function* m from A to a set M of objects called *meanings*. Now the core idea of compositionality, that the meaning of a compund expression is determined by the meanings of its parts and the mode of composition, is expressed as follows.

2.1 Definition. Let F be a k-ary operation of $\mathbf{A}$. m is *F-compositional* if there is a k-ary partial function G on M such that whenever $F(a_1, \ldots, a_k)$ is defined,

$$m(F(a_1, \ldots, a_k)) = G(m(a_1), \ldots, m(a_k)).$$

With G as above we say that m is *F-compositional with G*, and we say that m is *compositional* if it is F-compositional for all operations F of $\mathbf{A}$.

Note that we do not assume a semantic algebra given in advance, although if m is compositional, it induces such an algebra. If we make the further (reasonable) requirement that

(1) $$\text{if } m(a_i) = m(b_i), 1 \leq i \leq k, \text{ then} \\ F(a_1, \ldots, a_k) \text{ is defined iff } F(b_1, \ldots, b_k) \text{ is defined}$$

we can easily verify the following

2.2 Fact. *m is F-compositional iff whenever $m(a_i) = m(b_i), 1 \leq i \leq k$, and $F(a_1, \ldots, a_k)$ is defined, we have*

$$m(F(a_1, \ldots, a_k)) = m(F(b_1, \ldots, b_k)).$$

The following corollary is immediate:

2.3 Fact. *If m is one-one, then ((1) holds and) m is compositional.*

Usually, of course, m is not at all one-one: indeed one expects of a reasonable account of meaning that there are distinct expressions with the same meaning.

Fact 2.2 tells us what a claim of *non*-compositionality should look like: exhibit two expressions with the same meaning that nevertheless yield results with different meanings when a particular syntactic operation is applied to them. A familiar example is what Pelletier [8] dubs 'the argument from synonymy': let F be a 1-place operation which, when applied to a sentence φ, yields the sentence "John believes that φ". The claim is that any two synonymous expressions can be put in synonymous sentences φ and ψ such that "John believes that φ" and "John believes that ψ"may have different truth values, hence do have different meanings.

Another classical example is from Groenendijk and Stokhof [4] against a truth conditional account of meaning. Consider the two discourses

(2) A man is walking in the park. He whistles.

(3) It isn't the case that every man isn't walking in the park. He whistles.

Think of these as a syntactic operator H applied to sentences: $H(\varphi_1, \psi)$ and $H(\varphi_2, \psi)$. By truth conditional standards, φ_1 and φ_2 are synonymous. But $H(\varphi_1, \psi)$ and $H(\varphi_2, \psi)$ clearly differ in meaning, so, by Fact 2.2, we have a breach of compositionality.

Groenendijk and Stokhof regain compositionality by abandoning truth conditional semantics in favour of a more dynamic one. Others have tried (with less clear success) to counter the apparent non-compositionality of belief contexts. The question here is: Can we, in cases like these where compositionality is questioned, take relief in some general mathematical fact that a compositional semantics is *always* available?

3 Janssen's theorem

In [6], Janssen proved a theorem on how to make any meaning function compositional, a result which has been taken to show that compositionality has no empirical content. For example, Hendriks [5] writes: "...Janssen shows that any recursively enumerable language can be generated by an algebraic grammar. Moreover, such a language can be assigned any set of meanings in a compositional way. ...It follows that compositionality is not an empirical principle, but a methodological one." (p. 137).

In the present framework Janssen's theorem can be formulated as follows:

3.1 Theorem. (Janssen) *Suppose $m : L \longrightarrow M$, where L is any recursively enumerable set of strings and M is arbitrary. Then there is a partial algebra $\mathbf{A} = (A, F_0, \ldots, F_k)$ with $L \subseteq A$, a partial algebra $\mathbf{B} = (B, G_0, \ldots, G_k)$ with $M \subseteq B$, and a function h from A onto B which is F_i-compositional with G_i for $0 \leq i \leq k$, and such that for all $a \in L$, $h(a) = m(a)$.*

Pleasant as this result is, it is of no use to the semanticist. The crucial thing to notice is that *no structure of L is respected*; the theorem *existentially quantifies* over the structure (i.e., $\mathbf{A}$). Moreover, even if L is generated by some natural grammar, this grammatical structure is not reflected in the partial algebra $\mathbf{A}$ constructed in the proof. But clearly, without any *given* structure there just is no compositionality issue. Hence, Theorem 3.1 gives no reason to believe that compositionality is vacuous or non-empirical.

Let us note that Janssen himself, although he too regards the compositionality principle as methodological and not empirical, agrees that this does *not* follow from results such as Theorem 3.1 (cf. [7], p. 457). Indeed, with arguments similar to the one above he concludes that such formal results on the existence of compositional semantics are of no help in actually obtaining interesting semantic theories. (I believe the methodological/empirical distinction in the context of formal semantics warrants further elucidation, but in this note I am only concerned with the significance of the mathematical arguments.)

4 Free Algebras

In [3], Johan van Benthem sums up Janssen's work on compositionality in [6] like this: "The general outcome may be stated roughly as 'anything goes' — even though adherence to the principle [of compositionality] often makes for elegance and uniformity of presentation" (p. 57). He goes on to present the following mathematical argument for such a conclusion: "... the syntactic algebra ... is *free* (it is freely generated by the basic lexical elements). What this algebraic assertion amounts to is this. ... given any connection of [syntactic] operations ... with semantic operations of the same number of arguments, an arbitrary map from basic lexical items to suitable semantic entities will be uniquely extendable to a homomorphism as required. Thus we are entitled to conclude that *by itself, compositionality provides no significant constraint upon semantic theory.*" (ibid).

In this case syntactic structure *is* respected, so initially the argument looks more promising. The assumption is that the syntactic algebra is free, in fact, we can think of it as the (possibly partial) *term algebra* generated from the set of lexical items by the syntactic operators. This is indeed the usual situation in actual cases. Now suppose we fix, arbitrarily, the meanings of the lexical items. That is, the function m is defined for them. Also suppose we fix, again arbitrarily, the semantic operations in M which are to correspond to the syntactic operations. But then it is a fact of (universal) algebra that m can be uniquely extended to a homomorphism from our syntactic term algebra to the corresponding semantic algebra.

Doesn't this prove van Benthem's conclusion that compositionality by itself is no significant constraint? Not really. For the result presupposes a situation where *only* the meanings of lexical items have been fixed in advance. But this is not the case in serious arguments about compositionality. Both Pelletier's argument from synonymy and Groenendijk and Stokhof's argument against the compositionality of discourse when meanings are truth conditions, rely on definite intuitions about the meanings of certain *complex* expressions as well. And the extension of m to a homomorphism need not respect prescribed meanings of complex expressions. Indeed, if these arguments are correct, it *cannot* do so. In the Groenendijk and Stokhof case, for example (treating the subsentences as lexical items for simplicity), we have $m(\varphi_1) = m(\varphi_2)$ but are quite convinced that $H(\varphi_1, \psi)$ and $H(\varphi_2, \psi)$ differ in meaning. That is, however m is extended to a meaning function m^* defined on all expressions, we want $m^*(H(\varphi_1, \psi)) \neq m^*(H(\varphi_2, \psi))$. So m^* can never be a homomorphism, however the semantic operations are chosen.

Thus, relying on the above algebraic 'method' to enforce compositionality would mean giving up the intuition that (2) and (3) do not mean the same thing. Even if we recognize that linguistic intuitions are theory-dependent and need careful scrutiny (something which van Benthem stresses in [3]), this particular intuition looks like one we want to keep. And indeed it *is* compatible with a compositional semantics, namely, a dynamic one. But that semantics is hardly one we would get from a general mathematical existence theorem.

5 Concatenation as function application

Let us now look at the most recent theorem offered as an argument for the emptiness of the compositionality requirement.

5.1 Theorem. (Zadrozny [9]) *Let* $\mathbf{A} = (A, \cdot)$ *be a partial algebra, where the binary operation* $\cdot$ *can be thought of as concatenation (then* A *is a set of strings of symbols from some alphabet), and let* $m : A \longrightarrow M$*. Then there exists a set (of functions)* M^* *and a function* $\mu : A \longrightarrow M^*$ *such that for all* $a, b \in A$

(i) $\mu(a \cdot b) = \mu(a)(\mu(b))$ *(whenever defined),*

(ii) $\mu(a)(a) = m(a)$.

The step from M to M^* is described as a *type raising* by Zadrozny, and the moral of the result is that, by a suitable type raising, we can always obtain compositionality with *function application* as the semantic operation, and with the old meanings readily retrievable from the new ones.

Does this result prove the intended point about compositionality? And what is the role, and the justification, of the use of hypersets here?

On the latter issue, Zadrozny just says that "This set theory, ZFA [ZF set theory without Foundation but with Aczel's axiom of Antifoundation], is equiconsistent with the standard system of ZFC, thus the theorem does not assume anything more than what is needed for "standard mathematical practice". Furthermore, ZFA is better suited as foundations for semantics of natural language than ZFC (Barwise and Etchemendy, 1987)" ([9], p. 331). But this is rather misleading. First, hypersets have hardly become standard practice (yet). Second, mere equiconsistency with ordinary set theory is not a very good justification of their use. There are excellent such justifications, since they provide elegant models of various circular phenomena; cf. Barwise and Etchemendy [1], Barwise and Moss [2]. But, third, there seems to be nothing intuitively circular about the present situation. The 'type raising' in question, as Zadrozny himself stresses, is only a matter of enumerating all possible cases of combination, so that $\mu(a)$ is defined to be a function which takes as arguments all $\mu(b)$ such that a combines with b, giving the value $\mu(a \cdot b)$ (in addition it takes a itself as an argument so that $m(a)$ can be recovered).

Anything that can be described using hypersets can in principle also be described using only wellfounded sets. This follows from Aczel's consistency proof for ZFA, which builds a ZFA model $\mathbf{M}_A$ inside a standard model $\mathbf{M}$. Instead of using hypersets we could use the objects of $\mathbf{M}_A$, which are wellfounded from the point of view of $\mathbf{M}$, for our description. The price to pay is that the description may look complicated, since the epsilon relation in $\mathbf{M}_A$ is non-standard (it is not the restriction of the epsilon relation in $\mathbf{M}$ to $\mathbf{M}_A$).

For example, consider the set of all accessible pointed graphs[1] which are such that every node has a successor. This is an object of $\mathbf{M}_A$ and it can be seen that the only thing that stands in the epsilon relation (defined in $\mathbf{M}_A$) to this object is the object itself. From the perspective of $\mathbf{M}$, the object is a bit complicated, but if we allow hypersets, it is the simplest of all non-wellfounded sets, $\Omega = \{\Omega\}$.

Thus, although ordinary sets can always be used for semantic modeling, hypersets sometimes give a more elegant model. *If* we insist that function application should correspond to our binary syntactic operation,

$$\mu(a \cdot b) = \mathrm{APP}(\mu(a), \mu(b)),$$

hypersets elegantly allow this. Still, almost the same enumeration idea can be realized with ordinary sets. Given the assumptions of Theorem 5.1, define, for example,[2]

$$\nu(a) = \{\langle\langle a, 0\rangle, m(a)\rangle\} \cup \{\langle\langle b, 1\rangle, m(a \cdot b)\rangle \mid b, a \cdot b \in A\}.$$

Now $\nu(a)(\langle a, 0\rangle) = m(a)$ and $\nu(a)(\langle b, 1\rangle) = m(a \cdot b)$, so $m(a)$ can be recovered from $\nu(a)$, and it is rather clear that there exists an operation APP*, in some ways 'similar' to function application, such that

$$\nu(a \cdot b) = \mathrm{APP}^*(\nu(a), \nu(b)).$$

But why insist that function application is the only semantic operation? This is not required by compositionality, nor as far as I can see by other semantic considerations. It seems there is no other motivation than mathematical elegance. But this looks like elegance for its own sake: the crucial properties of the mathematical model have no interesting correspondence with properties of the thing being modeled (in contrast with the case when hypersets are used to model inherently circular phenomena).

So much for the use of hypersets. My main objection against Zadrozny's claim, however, is, as in the other two cases, that it gives no real help to the semanticist. The problematic issue is just avoided with an ad hoc construction.

[1] An accessible pointed graph $(G, \longrightarrow, p)$ consists of a set of nodes G, a set $\longrightarrow$ of edges (pairs of nodes), and a distinguished node p with the property that every node can be reached by some finite path from p.

[2] This kind of definition was suggested by Larry Moss.

Suppose we have found a counter-instance to the compositionality of m, say, expressions a_1, a_2 and b such that $m(a_1) = m(a_2)$ but $m(a_1 \cdot b) \neq m(a_2 \cdot b)$. We can think of Zadrozny's μ as giving a new meaning to a_1 as a huge list of ordered pairs, starting with $\langle a_1, m(a_1)\rangle$, and containing also $\langle \mu(b), \mu(a_1 \cdot b)\rangle$. Similarly, the list for a_2 starts with $\langle a_2, m(a_2)\rangle$ and contains $\langle \mu(b), \mu(a_2 \cdot b)\rangle$. Thus, $\mu(a_1) \neq \mu(a_2)$, and so the compositionality problem has just disappeared. But it disappeared for no interesting reason. No reason has been given why we should now say that a_1 and a_2 have *different* meanings, except for allowing ourselves the doubtful luxury of having $\mu(a \cdot b) = \mu(a)(\mu(b))$.

6 Conclusion

The three results we have mentioned achieve compositionality by syntactic and semantic manipulations that are of no interest to the semanticist. Janssen's theorem introduces a completely ad hoc syntax. Automatically extending a mapping of lexical items to a homomorphism as van Benthem suggests violates basic intuitions in actual examples. Zadrozny's theorem, the most striking result in this vein, makes the meaning assignment one-one in an unmotivated way, thereby side-stepping the compositionality issue.

Furthermore, if ad hoc moves are to be allowed, compositionality can be achieved much simpler than in any of the three mentioned results:

6.1 Fact. *Suppose* $\mathbf{A}$ *is a partial algebra and* $m : A \longrightarrow M$. *Let* $M' = A \times M$. *Then* $m' : A \longrightarrow M'$ *defined by*

$$m'(a) = \langle a, m(a)\rangle$$

is compositional.

Proof. By Fact 2.3, since m' is one-one.

Here too m is readily recoverable from m', by $m(a) = 2^{nd}(m'(a))$. But the point is that anyone can see that Fact 6.1 is completely trivial, and, hopefully, that it therefore can have nothing of interest to tell us about the status of compositionality. My claim, then, is that the three technical results considered here have *as little* to contribute in this respect as Fact 6.1.

It might be objected that I am making more out of these three cases than the respective authors intended. After all, Janssen explicitly warns against misusing his theorem, and surely van Benthem would not recommend practical use of the algebraic 'method' as I suggested in section 4. Even Zadrozny, in the quote from the Introduction, only claims that compositionality is "formally" vacuous, and in fact he is aware in that paper that a semanticist will want further constraints.

But, with such disclaimers, notice how weak the dictum that any semantics can be made compositional becomes. In Zadrozny's case (disregarding the —

doubtful — part about concatenation and function application), 'Any semantics can be made compositional' amounts to the claim that for any meaning function m from a partial algebra $\mathbf{A}$ to a set M, there is *another* meaning function μ from $\mathbf{A}$ to *another* set M^* such that (a) μ is compositional, and (b) $m(a)$ can be recovered uniformly from $\mu(a)$. Similarly for the other two cases. Now, stated thus explicitly, this claim is *trivially* true (cf. Fact 6.1), and the mathematical machinery invoked in the three examples is quite unnecessary. But then why invoke it? Surely it is to give us the impression that these pieces of mathematics provide some useful or interesting methodological insight into the nature of the principle of compositionality. That impression is precisely what I have tried to dispel in this note.

References

[1] J. Barwise and J. Etchemendy, *The Liar: An Essay on Truth and Circularity*, Oxford UP, 1987.

[2] J. Barwise and L. Moss, *Vicious Circles*, CSLI Publications, Stanford, 1996.

[3] J. van Benthem, The logic of semantics, in *Essays on Logical Semantics*, D. Reidel, Dordrecht, 1986, 198–214.

[4] J. Groenendijk and M. Stokhof, Dynamic Predicate Logic, *Linguistics and Philosophy* 14 (1991), 30–100.

[5] H. Hendriks, *Studied Flexibility. Categories and Types in Syntax and Semantics*, ICCL Dissertation Series 1993–5, University of Amsterdam.

[6] T. Janssen, *Foundations and Applications of Montague Grammar, Part 1*, CWI Tract no. 19, Center for Mathematics and Computer Science, Amsterdam, 1986.

[7] T. Janssen, Compositionality, in J. van Benthem and A. ter Meulen (eds), *Handbook of Logic and Language*, Elsevier, Amsterdam, 1996, 417–473.

[8] F. J. Pelletier, The principle of semantic compositionality, *Topoi* 13, 1994, 11–24.

[9] W. Zadrozny, From compositional to systematic semantics, *Linguistics and Philosophy* 17 (1994), 329–342.