

Is the Compositionality Principle a Semantic Universal?

Author(s): Cynthia K. Welsh

Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society (1986), pp. 551-563

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

Is the Compositionality Principle a Semantic Universal?*

Cynthia K. Welsh
University of Chicago

1. Introduction

A passage from the gospel of semantics (if such a thing existed) might read:

The meaning of the sentence must be such that it is a function of the meaning of its component parts and the way in which they are combined. This is the principle of compositionality, and it is a fundamental constraint on semantic theories. And I say unto you, whatsoever is not compatible with this principle, is not within the domain of semantics.

This is the gospel according to Frege.

While putting it this way might seem a little silly (we are, after all, empirical investigators, not religious fanatics), the fact is that compositionality (CP), or the Fregean principle is often adhered to and defended with a faith and fervor quite like that exhibited by "believers" defending religious doctrine. Why is this?

On the one hand, there are good reasons to maintain the CP as a constraint on semantic theory. First, the empirical evidence for so constraining a semantic theory is based on the fact that speakers can figure out the meaning of a complex expression they've never before encountered based solely on the meanings of the constituent expressions and the mode of syntactic combination. This evidence seems to indicate that compositionality is a universal of natural language semantics, and therefore justifies such a constraint. Second, there is the sort of what-else position - if semantic analysis isn't compositional then what else can it be. Compositionality, from this point of view, is simply what we mean by semantic analysis. Therefore the CP is true by definition, so to speak. The CP, from either point of view, appears to be a well-motivated universal constraint on NL semantics.

On the other hand, dogmatic defense of the compositionality principle, especially under the interpretation that CP-Montague's theory of grammar, has led to a lot of confusion. Recently in the literature, for instance, psychologically plausible alternatives to the classical theory of lexical semantic representation, such as certain instantiations of prototype theory, are deemed inadequate because

they are not (and cannot be) compositional (see for instance, Osherson and Smith, 1981; Armstrong, Gleitman and Gleitman, 1982). The general argument for this point of view can be summarized as in (1):

- (1) a. An adequate semantic theory must be compositional.
- b. The only possible combinatorics for prototype theories is some version of fuzzy-set theory.
- c. Fuzzy-set theory yields the wrong results for conjunctive concepts, inclusions and related matters.
- d. Therefore, prototype theory is not (and, some would argue, cannot be) compositional.
- e. Therefore prototype theory is inadequate as a semantic theory.

There are some flaws and unsupported assumptions in this argument which I won't go into here. What I am most concerned with is the range of responses that have been made to this argument. One counter to this argument develops an alternative version of fuzzy-set theory, which is claimed to be compositional under a broader construal of the notion of compositionality. For instance, Zadeh (1983) claims:

In its traditional interpretation, Frege's principle of compositionality is not sufficiently flexible to have a wide applicability to natural languages. In a fuzzy-set-theoretic setting which is outlined in this paper, Frege's principle is modified and broadened by allowing the meaning of a proposition P to be composed not from the meaning of the constituents of p, but more generally, from the meaning of a collection of fuzzy relations which form a so-called explanatory database that is associated with p.

Another response is to develop a model of prototype representations which are modified by principles of conceptual combination to generate complex representations which yield the correct typicality results, and are therefore claimed to obey compositionality. Smith and Osherson (ms) use this tactic. Finally, others have denied that our complex representations obey compositionality, thereby denying that the principle characterizes a semantic universal. Cohen and Murphy (1984) take this position. The question I am here to address is whether compositionality is a semantic universal, and why, when addressing the same empirical evidence, are modern theoreticians led to such divergent claims regarding the status of compositionality.

2. Versions of Compositionality

I contend that the problem stems from a confusion about what "version" of the compositionality principle is under discussion. We all seem to assume there is simply one unified notion of the compositionality principle. The use of the common appellation across theories and the general lack of detail about what the principle entails, seems to justify the validity of such an assumption. On closer review, it becomes apparent that compositionality principles range from strong versions, such as that employed in Montague's Universal Grammar, to versions so vague it is difficult to see how they would impose any constraint at all on semantic theory. In its most general form the principle can be stated as in (2):

- (2) The meaning of a complex expression is a function of the meaning of its parts and the way in which they are syntactically combined.

Much more needs to be specified for this general form of the principle to bear any weight - for example we need to specify what meanings and functions are and what can count as the relevant component parts. It becomes apparent when considering these questions that the principle, theory-independently, has little empirical content. It is only with the support of frequently tacit, theory-relative assumptions that the principle has any teeth. There are (at least) four supporting assumptions which characterize the strictest version of compositionality à la Montague, (1970). Simplifying considerably, and avoiding technicalities, these assumptions can be characterized as in (3) - (6):¹

(3) homomorphism - requires that there is a parallelism between the syntax and semantics, such that for each syntactic entity there is a corresponding semantic entity. More specifically, if any expression of syntactic category x corresponds to the semantic type x , then all expressions of category x must correspond to semantic type x .

(4) locality - requires that interpretations are built bottom-up and prohibits global properties of the complex expression from affecting the semantic value of constituents.

(5) meaning invariability - requires that there is one and only one contribution made by an expression e to the meaning of any complex expression E in which e occurs. This contribution cannot vary from context to context.

(6) determinacy - requires that the meaning of a complex expression E must be completely determined by the constituent expressions $e_1 \dots e_i \dots e_n$ of which it is composed. That is, any aspect of meaning with

which the complex expression is endowed, must be traceable to one of the constituent elements, or to the construction itself.

While these assumptions are clearly interrelated and seem to form a coherent package, one can adhere to or abandon one or another of these assumptions without committing oneself to buying or abandoning the whole package. Even abandoning all of the assumptions in (3) - (6) does not necessarily commit one to abandoning the spirit of the compositionality principle: that the assignment of meanings to complex expressions is a *systematic* (i.e. predictable or recursive) function based on the meaning of the parts.

I shall have little more to say directly about assumptions (4) and (6). Assumption (5) seems untenable within a prototype-like theory, because we would presumably want to say that contributions can vary without having to posit ambiguity. In fact, I think it is the desire to maintain assumption (5) that has led some theorists to posit hybrid theories of "core + other" criticized by Lakoff (1982 ms).

In the balance of this paper I will provide evidence that assumption (3) is too strong a constraint for natural language semantics. Finally, I will show how a prototype-like theory can conform to a modified compositionality constraint, which abandons assumptions (3) and (5).

3. Challenges to Homomorphism

The assumption of homomorphism is a cornerstone of Montague semantics and is often simply identified with the compositionality principle. One interesting consequence of assuming homomorphism, noted by Partee (1984), and Landman and Moerdijk (1983), is that it makes an intermediate level (such as the logical language of translation, or discourse representations a la Kamp) dispensable in principle. Since the compositionality principle requires that for any translation rule:

(7) $F_i(a,b)$ translates as $G_j(a'b')$

the operation G_j must be such that for semantically equivalent input it yields the same results. The only properties which can play a role at the intermediate level are properties which are already reflected at the syntactic and/or semantic levels, since under a strict

interpretation of (3), one level simply recapitulates the other. Thus, in principle, an intermediate level must be dispensable.

One sort of challenge to a strict homomorphic constraint, comes from arguments and analyses which require an indispensable level of representation not homomorphically related to the syntactic and semantic levels. An example of this sort of broad challenge to (3) is exhibited recently by the work of Kamp (1981) and Heim (1982). These theories, using an indispensable level of representation, provide a type of solution to the puzzle of donkey-sentences superior to other analyses. One can argue on the basis of examples like these, that since positing an indispensable level of representation allows for an elegant analysis of some natural language phenomena not heretofore adequately treated, then the homomorphic constraint, which disallows such a level, must be too strong.

Another sort of challenge concerns the correspondence required between syntactic and semantic categories. Williams (1983) presents evidence that logical distinctions do not always correspond to syntactic distinctions in his discussion of predicate and referential nominals. He argues that since both types of nominals have the same internal syntax - that is both conform to the PS rule (8):

(8) NP \rightarrow det AP N PP S

and both allow free relatives, then both types of nominals should be assigned to the same syntactic category. But, in spite of their belonging to the same syntactic category, he contends they differ in logical type. Predicate nominals, unlike referential nominals translate into one-place predicates, as is illustrated by (9) - (11):

(9) John became a doctor.

(10) * A doctor was become by John.

(11) A doctor was seen by John.

Predicative nominal phrases, like one-place predicates, must have a subject. In (9) 'John' is the subject of 'a doctor'. (10) is unacceptable because the predicate nominal is not c-commanded by its subject. Referential NP's, on the other hand, do not have subjects - they are not functors which require arguments and therefore (11) is acceptable. The explanation of the difference between (10) and (11), Williams argues, rests on positing a distinction in logical type. Since

this logical distinction is not matched by a syntactic distinction, this is a violation of assumption (3).

Another important example that challenges assumption (3), discussed most recently by Higginbotham (1985), concerns the interpretation of non-well-formed-formulas. Consider (12):

(12) * The child seems sleeping.

(12) can only be understood in the same way as "the child seems to be sleeping"; the grammar somehow disallows the interpretation of the gerund "sleeping" as an adjective as in "The child seems sleepy."

However, since (12) is ill-formed, presumably there would be no syntactic rule to generate it. Given the strict homomorphic assumption in which syntactic entities are paralleled by semantic entities, it seems, then, that there will be no corresponding semantic rule to interpret it, let alone predict its definite interpretation.

As a final challenge, consider what one must commit to if one assumes both (3) and (6) when analyzing constructions with apparent extra meanings or "missing" but understood elements. Consider (13):

(13) Being a master of disguise, Bill would fool anyone.

Stump's (1985) analysis of these sorts of constructions involves the introduction of a free variable which ranges over propositional relations. If one is firmly committed to assumptions (3) and (6) together, one is led to analyses of these sorts of constructions that enrich, or "build in" elements at the syntactic and/or semantic levels so that (3) and (6) can be upheld. This is one option in the analysis of the above, and this strategy has often been adopted for constructions which have "missing" (syntactically) but understood arguments. This strategy for maintaining assumptions - the brute force strategy - seems particularly unappealing and counterintuitive. This is especially so when one considers that there is a plausible alternative: to abandon homomorphism. This alternative allows you to maintain determinacy, which is then defined over the intermediate level of representation, and obviates the need to build-in things that don't seem to be there syntactically - things to which the syntax appears to be indifferent.

On the basis of this type of evidence, I contend that the homomorphism constraint is too strong for natural language. One

could of course, do patch-work repairs to save the assumption but this seems rather unintuitive and doesn't seem to buy you anything.

Abandoning assumptions (3) and (5) allows us to maintain a modified version of the CP which can be characterized as follows:

First, interpretations are built up locally, therefore meeting the locality requirement.² Second, by abandoning (3) we are able to introduce an intermediate level of representation. It is at this intermediate level that the relevant parts are found, and therefore it is to this level that CP applies. Further, the combinatorial operations are over structured representations at this intermediate level. The operations are not restricted to the basic operations on sets. Rather representations can be altered in various ways in the combinatoric process, as will be shown in § 4 and § 5 below.

4. The representations

To represent the core of a lexical item, I will adopt Moravcsik's (1981) aitiational frame format. The aitiational frame is construed as an explanatory schema, which consists of four meaning parameters and a conceptual type.³ The four meaning parameters can roughly be characterized as in (14):

- (14) m-parameter: the kind, the stuff it's made of, its essential parts
s-parameter: principles of individuation distinguishing it within the kind.
f-parameter: the characteristic associated function
a-parameter: the characteristic causal or interactional properties.

What Rosch's results have taught us is that there are no necessary and sufficient conditions for most of our vocabulary. Rather, there are many cases in which a single lexical item seems not to correspond to a single, well-defined or bounded concept, but seems rather to be related to systematically connected families of concepts. The lexical specifications provided in the aitiational frame form a core around which an interrelated family of concepts are generated. Each family 'member' may, under appropriate interpretational circumstances, contribute to or become the interpretation, in accord with general combinatoric principles

The specifications provided by the aitiational frame are not necessary and sufficient conditions, nor does the frame provide more than partial specifications. Rather, the frames are to be construed as providing minimal "default" specifications. That is, these are the feature

specifications incorporated in the interpretation of a lexical item unless explicitly defeated, or unless some conflicting alternative specification is provided. As an example consider the specification of lemon in (15):

(15) lemon (CN)

<u>type</u>	<u>m-factor</u>	<u>s-factor</u>	<u>f-factor</u>	<u>a-factor</u>
Nk	fruit	yellow	edible	...
object		round		
		sour		

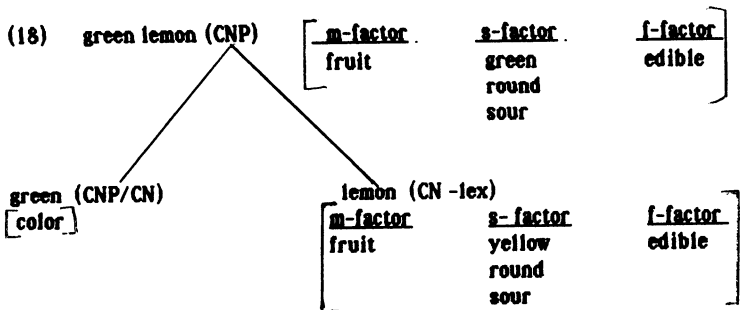
The m-factor specifies the basic domain: 'fruit', with the s-factor specifications distinguishing it from other objects falling in the same domain; e.g. the specification 'yellow' distinguishes lemons from limes, the specifications 'round' and 'sour' distinguish it from bananas and yellow apples. Our typical interpretation of 'lemon' incorporates these 'default' features, as in (16), but we certainly understand (17), even though one of the default values has been defeated.

(16) I bought a lemon yesterday.

(17) I bought a green lemon yesterday.

We would not want to say that the typical occurrence of 'lemon' (as in (16)) and its occurrence in (17) mean different things. Rather, one of the default values has been defeated in (17) by the specification of a conflicting alternative: 'green'.

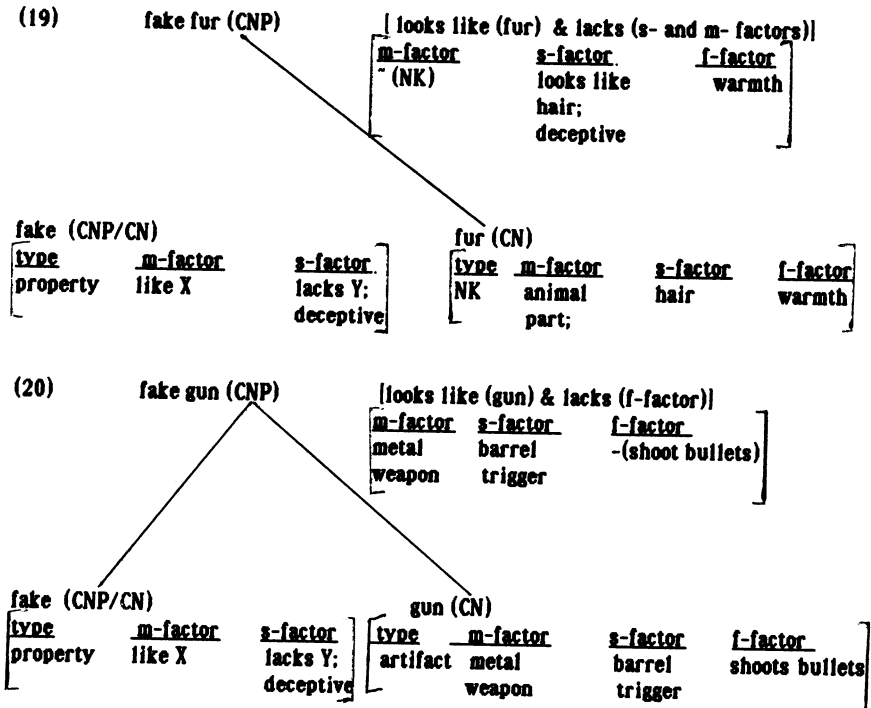
An alternative to the default value replaces the default value, it is not simply added to it as illustrated in (18):



Another feature of these representations is that the type and the m-factor determine the dominant (or important) parameter(s) creating

a hierarchical structure for lexical representations. Artifact terms, are generally dominant along the f-factor, followed by their related causal roles. Natural kind terms, on the other hand, are dominant along the structure parameter. As an illustration of this thesis, consider phrases like 'fake CN':

Intuitively we interpret this type of construction to mean something that resembles the CN in appearance, but lacks certain other features characteristic of the real CN. What these characteristics are, is determined by the dominant parameter of the CN. Consider the phrases "fake fur" and "fake gun" in (19) and (20):⁴



Fake is a function which takes a CN as an argument, and defeats the dominant parameter determined by the conceptual type of the CN. The value of the resultant complex expression is a representation that consists of all but the features of the dominant parameter of the argument. Applying this function to 'fur' and 'gun' as arguments, we find intuitively, that a fake fur must share the features which serve to define a real fur, save its structure. It must look like a fur and provide

warmth, but it would lack the characteristic structure (including origin) of real furs. A fake gun, on the other hand, must share the features which define a real gun, save its function and related causal power. A fake gun looks like a gun but it doesn't shoot bullets.

The hierarchy of parameters gives us a way to capture these intuitions. The difference between 'fake' when applied to 'gun' and when applied to 'fur' can be attributed to the difference in the type and m-factor in the two terms, which determine the dominant parameters. 'Fake' functions to defeat just these dominant parameters.

5. The Combinatorics

In addition to the notion of defeat introduced above and the standard combinatoric principles of conjunction, disjunction and negation, operational in the traditional semantic theory, I claim there are (at least) two other combinatoric principles operational in the interpretation of complex expressions. Consider the differing interpretations of "letter" in (21)-(24) (examples from Bierwisch 1982):

- (21) I put the letter on your desk.
- (22) The letter has been distributed to the whole faculty.
- (23) The letter finally led to a political crisis.
- (24) For many poets, the letter is a genuine literary genre.

The basic specification for letter would be as in (25):

(25) letter (CN)				
<u>type</u>	<u>m-factor</u>	<u>s-factor</u>	<u>f-factor</u>	<u>a-factor</u>
object	written x	y address x	information	...
		to z		

In (21) the most natural interpretation of Letter, is as a physical object of a certain kind. This interpretation is in line with the basic frame in (25). In (22) the interpretation has shifted types, from basic object, to a set of those objects. In (23) the type has shifted to the informational content, and finally (24) has shifted to the type of informational structure.

In these examples, our interpretation has shifted from the basic concept to focus on one of the related members of the complex concept. Following Bierwisch, I will call the principle governing this sort of shift in interpretation "conceptual shift".

Another sort of principle involved in alternative interpretations is "conceptual specification". Consider the situational specification for 'lose' in (26):

(26) lose (TV)

<u>type</u>	<u>m-factor</u>	<u>s-factor</u>	<u>f-factor</u>	<u>a-factor</u>
event z	cause to not	x have y at t &	...	x not have y
	have y	not have at t' >		

Now consider the different ways of losing involved in (27) - (31):

- (27) John lost his money, as he had a hole in his pocket.
- (28) John lost his money by speculating at the stock market.
- (29) John lost his friend in the crowded train station.
- (30) John lost his friend in a tragic car accident.
- (31) John lost his friend, as he always told bad jokes about him.

In (27) the event is further specified as a change in location of a concrete object 'money'. In (28) the event is specified as a change in possession of an abstract exchange value. In (29) the event is a loss of knowledge of location of a person. (30) seems to refer to an event of loss of possession of a relation to a person (as a result of a change in that person's state). And (31) it is the loss of abstract possession of a 'friendship'. Each of these differing interpretations are the result of specifications of values of the event variable induced by the context.

Conceptual specification is different from conceptual shift. Conceptual shift involves the shift from categories of concepts of one type, to concepts of another related type. Conceptual specification, on the other hand, provides different specifications of variables, resulting in different concepts within one and the same conceptual type. I contend that you need at least these two combinatoric principles, in addition to the traditional operations of conjunction, disjunction and negation, in order to account for complex categorization.

6. Conclusion

What I have attempted to illustrate in this paper is that prototype-like representations can be incorporated in a compositional semantics. I have argued that there is good reason to abandon the homomorphic assumption, and that doing so allows an intermediate level of representation to which the compositionality principle can apply. These representations need to be structured and the combinatoric operations which operate on these representations must be able to "look inside" the representations, and modify them in accord

with the principles of defeat, shift and specification, when they occur as part of a complex expression. Finally, this way of construing prototype representations and compositionality does not violate the spirit of compositionality, even though it does abandon some of the typical supporting assumptions. The complex representations are constructed in a predictable and rule governed way from the constituent representations. I believe it is this systematicity which is at the heart of the compositionality principle taken as a semantic universal.

NOTES

* My thanks to those at BLS and in Chicago who have commented on various stages of this work. Special thanks go to R. Chametzky, D. Dowty, G. Lakoff, J. McCawley, R. McClamrock, J. Moravcsik, J. Richardson, J. Sadock, M. Silverstein and J. Stern, whose discussions, comments, and criticisms have helped me to correct errors and clarify my thinking. Remaining mistakes, of course, are my own.

1. Hintikka (1981) discusses some of these assumptions in a different context and uses slightly different terminology. I find my terminology more perspicuous.

2. Under the outline of a theory developed below, this might not always be the case, since representations can be altered and I have assumed nothing which would close representations to modifications "higher" up in the derivation. For the purposes of this paper, however, I will simply assume locality.

3. I believe this representational format has advantages over other options. First, it allows one to represent the grouping or constituency of values, rather than simply presenting unordered and unprincipled lists. Second, it allows one to characterize the dominance of certain types of values, giving a hierarchical structure to the representations. The conceptual type is my own innovation adapted from Bierwisch (1982). I'm not sure that either Moravcsik or Bierwisch would approve of my construal. Also, the representations provided below are rough characterizations for illustrative purposes only. I'm not religious about the details.

4. David Dowty pointed out to me, during the discussion period, that these particular examples could be handled within a framework that assumed homomorphism. I am grateful to him for pointing out that the point of these examples wasn't clear in the presentation. I am using these examples here only to illustrate the hierarchical character of the representations. Where I think this framework excels with respect to this data is first, it predicts what properties a "Fake CN" will and won't have, based on the frame for the CN, rather than just saying what they are not (that is, mapping them into some unspecified subset of the complement of CN). Second, since these are only partial representations, it is predicted that the fake CN will lack whatever values the dominant parameter contains, even if we have no specific information represented under that parameter. For example, we know that fake furs will lack the essential structural properties (and, perhaps causal history) of real furs, even though most of us are ignorant as to what these properties are or how to identify them. This seems just the right prediction. George Lakoff pointed out that hedges like "strictly speaking" and "technically" are better examples of expressions that can't be

handled within a homomorphic system. The examples provided under section 5. below, are examples of things that can't be handled within theories which assume homomorphism and which restrict themselves to the basic set operations.

REFERENCES

- Armstrong, S., L. Gleitman, and H. Gleitman. 1982. "What Some Concepts Might not Be". Cognition 13.
- Bierwisch, M. 1982. "Formal and Lexical Semantics". Linguistic Berichte 80/82.
- Cohen, B. and G. Murphy. 1984. "Models of Concepts". Cognitive Science 8.
- Heim, I. 1982. "The Semantics of Definite and Indefinite Noun Phrases". Ph.d thesis reproduced by Graduate Linguistic Student Association, University of Massachusetts.
- Higginbotham, J. ms. "Linguistic Theory and Davidson's Program in Semantics".
- Hintikka, J. 1981. "Theories of Truth and Learnable Languages". Philosophy and Grammar. S. Kanger and S. Ohman, eds.
- Kamp, J. 1981. "A Theory of Truth and Semantic Representation". Formal Methods in the Study of Language. J. Groenendijk, T. Janssen and M. Stokhof, eds.
- Lakoff, G. 1982 ms. "Categories and Cognitive Models".
- Landman, F. and I. Moerdijk. 1983. "Compositionality and the Analysis of Anaphora". Linguistics and Philosophy 6.
- Montague, R. 1970. "Universal Grammar". Formal Philosophy. R. Thomason, ed.
- Moravcsik, J. 1981. "How do Words Get their Meaning?". Journal of Philosophy. LXXVIII no.1.
- 1982. "Aitiational Frameworks and the Structure of Meaning". Acta Linguistica Academiae Scientiarum Hungaricae. 32.
- Osherson, D. and E. Smith. 1981. "On the Adequacy of Prototype Theory as a Theory of Concepts". Cognition 9.
- Partee, B. 1984. "Compositionality". Varieties of Formal Semantics. F. Landman and F. Veltman, eds.
- Smith, E. and D. Osherson. ms. "Conceptual Combination with Prototype Concepts".
- Stump, G. 1985. The Semantic Variability of Absolute Constructions.
- Williams, E. 1983. "Semantic vs. Syntactic Categories". Linguistics and Philosophy 6.
- Zadeh, L. 1983. "A Fuzzy-Set-Theoretic Approach to the Compositionality of Meaning: Propositions, Dispositions and Canonical Forms". Journal of Semantics 2.