

Is Compositionality an A Priori Principle?

Daniel Cohnitz

When reasons are given *for* compositionality, the arguments usually purport to establish compositionality in an almost a priori manner. I will rehearse these arguments why one could think that compositionality is a priori true, or almost a priori true, and will find all of them inconclusive. This, in itself, is no reason *against* compositionality, but a reason to try to establish or defend the principle on other than quasi-a priori grounds.

I want to argue in this paper that there is a *substantial* (non vacuous) notion of compositionality that seems to be of interest for semanticists. There is also an argument for this kind of compositionality that superficially looks like one from the standard battery of arguments for compositionality but that is not quasi-a priori in the same way as the others are. Instead it rests on an empirical hypothesis of which we do not know on quasi-a priori considerations alone whether it is true or false.

1 Introduction

A superficial look at the literature on the principle of compositionality (henceforth ‘The Principle’) could suggest that the discussion is as confused as a discussion can be. This starts already with the question of the proper historical origin of The Principle. Although it is often called ‘Frege’s Principle’, it is controversial whether Gottlob Frege subscribed to The Principle throughout his work, or even at certain stages of his intellectual development. The main reason for *that* controversy is that Frege is also famous for another principle, the so called ‘Context Principle’, or ‘Principle of Contextuality’ which is *prima facie* in tension with the idea of compositionality. Whereas compositionality seems to explain the meaning of linguistic expressions bottom-up, by saying (in one way or other) that the meaning of a complex expression is a function of the meaning of its subexpressions, the Principle of Contextuality explains the meaning of expressions top-down, by saying that it is the meaning of the whole in virtue

Address for correspondence: Department of Philosophy, Heinrich-Heine University Düsseldorf, D-40225 Düsseldorf, Germany.

E-mail: cohnitz@phil-fak.uni-duesseldorf.de.



A quick guide to different interpretations of “Frege’s Principles”

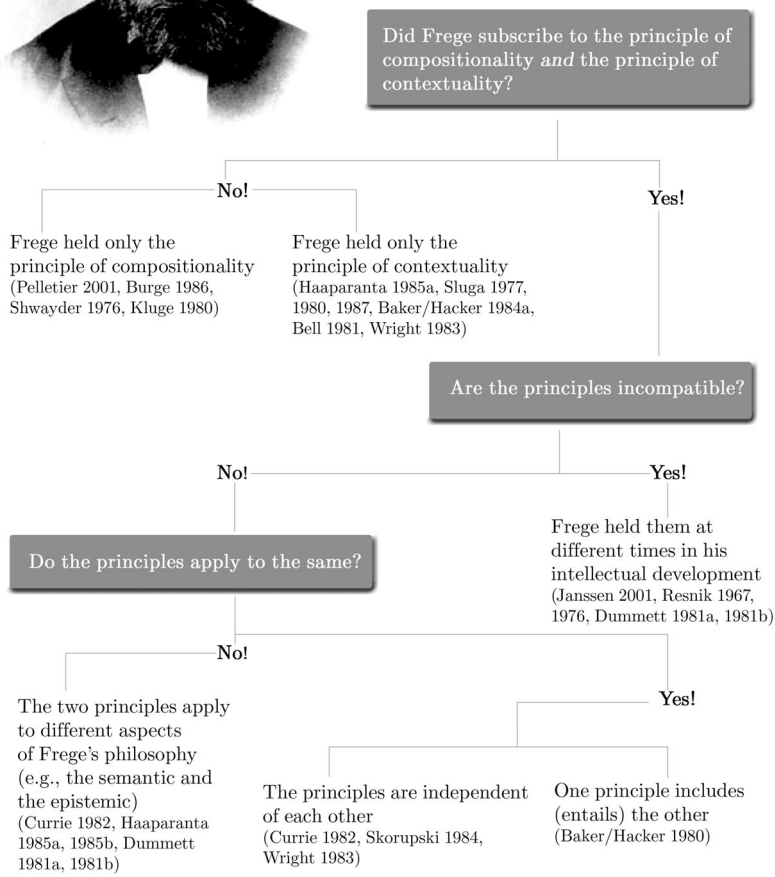


Figure 1: A guided tour to the different interpretations of Frege’s two principles.

of which the subexpressions have a determinate meaning. The subtleties of Frege exegesis are not our concern here, which is why figure 1 might suffice as a guide through the different positions that were defended with respect to this controversy.¹

However, a quick look at figure 1 already highlights a second difficulty of any discussion of The Principle: all parties that do not disagree with respect to the first question disagree in their interpretation of at least one of the principles. They, for example, disagree on whether the principles concern reference or meaning or both or on whether they are meant in an epistemological sense or rather ontologically. This suggests that the *prima facie* understanding of The Principle and the Principle of Contextuality are misleading.

Although the principles *seem* to be talking about the same and *seem* to be inconsistent, some argue that they not really *are* talking about the same or not really *are* inconsistent. Obviously there are different ways to understand The Principle, ways that are not only relevant for an adequate interpretation of Frege, but also ways that are relevant for a proper assessment of the epistemic status of The Principle. This leads us to the second aspect which is unclear about The Principle: what is it about and what exactly does it say?

Whereas we will exclusively consider The Principle as intended to apply to (spoken) natural languages, the principle of compositionality is often intended to be about the *mental representations* correlated with elements of natural languages and concerns their syntax and their semantics (if mental representations have these). Such mental representations might be – for example – prototype-structured and then might or might not be compositional themselves (see Machinery, 2001; Fodor, 1998; Fodor & Lepore, 2002). We will just note that the question of whether the ‘language’ of thought (LOT) is compositional might be different from the question of whether natural language is. For example, natural language might be non-compositional because of the fact that it contains synonymous sentences and the possibility to embed them in belief-contexts (Pelletier 1994), whereas it seems arguable, on the other hand, that LOT would not have two representations for the same meaning.² A falsification of composi-

¹For a guide to this guide and for page references see Pelletier (2001) where this classification was invented. The references indicate places where one can either find the interpretation in question or a reconstruction of it. Some authors consider alternative possible interpretations, which is why they might occur in positions in the tree which are actually inconsistent.

²Note, however, that this reply loses plausibility if beliefs are tokened sentences of LOT in a belief-box the way ‘Harvey’ is modeled in part (3.). In this case all synonymous expressions of the spoken language would correspond to one and the same sentence of LOT (that would be synonymy) and therefore embedding them into belief-contexts could not violate compositionality anymore (always the same sentence of LOT would

tionality of natural language (e.g., if based on the existence of synonymies) is therefore no straightforward falsification of the compositionality of LOT. However, the arguments we will consider that speak in favor of compositionality can sometimes be reformulated to speak in favor of the compositionality of LOT.

But even if we want to understand compositionality as applied to natural language, it is still unclear what The Principle actually says. Consider the standard³ formulation of The Principle:

PoC 1 (Principle of Compositionality). The meaning of an expression is a function of⁴ the meanings of its parts and of the way they are syntactically combined. (Partee, 2004, p. 153)

As Barbara Partee (2004, p. 154) has emphasized already, this principle can only be made precise in the context of a theory. According to Partee, The Principle is theory-dependent in at least these respects:

i. What are the ‘meanings’? Are they considered model-theoretic objects?, linguistic representations?, intensions?, functions from contexts to intensions?, etc.

ii. What is assumed about the syntax? Is it independently motivated? Is it constrained by compositionality? What kinds of abstractness and invisibilia are allowed?, etc.

iii. How is the ‘is a function of’ relation to be understood? Are there any constraints on what kinds of functions interpret what kinds of syntactic combinations? Is compositionality necessarily purely bottom-up? Must the functions be single-valued? Does functionality preclude non-dispensable intermediate levels of representation?, etc.

Given that The Principle is not very precise if these questions are not answered beforehand, it is understandable that the The Principle is often regarded to be an
 be tokened by synonymies and therefore substitution of synonymies in belief-contexts in the natural language would preserve truth (and meaning)).

³...if there is any such thing as ‘the’ standard formulation. Zoltán Szabo (2000a) found six formulations of compositionality in one and the same logic textbook only one of which he considered a stylistic variant of the principle of compositionality.

⁴As Pelletier has emphasized, the principle should rather read ‘... is a function of, and only of, ...’. Pelletier’s point is that in the reading given above it seems allowed that the meanings of the parts could only be dummy variables of the meaning-determining function that play no actual role in determining the meaning of the whole (which is instead determined by the ‘real’ arguments of the meaning-determining function). That is of course not intended by the friends of compositionality. What they intend to say is that the meanings of the parts are the real arguments determining the meaning of the whole, hence the ‘only of’-clause. See Pelletier (1994, p. 11).

obvious, almost self-evident truth. As we will see, rather plausible constraints on the possible answers to the questions above will weaken the standard arguments *for* compositionality considerably. But let's consider why one could think that The Principle is true a priori.

2 Formal Vacuity

If it could be shown that compositionality is formally vacuous, compositionality would of course be an a priori principle, but at the same time absolutely uninteresting.⁵ Just as

PoC 2. Human languages have a compositional semantics or human languages don't have a compositional semantics.

does not tell us much about the semantics of human languages, compositionality equally would not tell us much. Some mathematicians/logicians have claimed that compositionality is so vacuous. One example is Zadrozny:

We prove a theorem stating that any semantics can be encoded as a compositional semantics, which means that, essentially, the standard definition of compositionality is formally vacuous. (Zadrozny, 1994, p. 329)

Another example is van Benthem:

[Frege's principle of compositionality] has been investigated thoroughly in an algebraic setting in Janssen, 1983. The general outcome may be stated roughly as 'anything goes' – even though adherence to the principle often makes for elegance and uniformity of presentation. [...] Thus, by itself, *compositionality provides no significant constraint upon semantic theory*. (van Benthem, 1986, p. 200)

As we have said already, The Principle is taken to be the thesis that the meaning of a whole is a function of the meaning of its parts. If we state the principle that loosely, compositionality may not be very substantial. The results obtained by Janssen (1983), van Benthem (1986) and Zadrozny (1994) are sometimes interpreted to have shown that for every language the semantics for that language can be represented as a compositional semantics. What did they do?

Janssen was the first to prove that for any language and any meaning, the meaning can be assigned to the language in a compositional way. To get to this result, Janssen exploits ideas from Montague (1970) and the theory of universal

⁵Which is not very surprising for a good empiricist: all a priori truths (if there are any) are not going to be informative.

algebra. The key idea that led to a formal treatment of compositionality is that compositionality requires the existence of a *homomorphism* between the expressions of a language and their structure and the meanings of those expressions and their structure. The *relata* of this homomorphism are formally represented as *algebras*.

D 1 (Algebra). An algebra $\mathcal{A}$, consists of a set A called the *carrier* of the algebra, and a set F of functions (operators) of that set and yielding values in that set. So $\mathcal{A} = \langle A, F \rangle$. The elements of the carrier are called the *elements of the algebra*. If an operator (function) is not defined on the whole carrier, it is called a *partial operator*. If $E = F(E_1, E_2, \dots, E_n)$, then $E_1, E_2, \dots$, and E_n are called *parts of E* . If an operator takes n arguments, it is called an n -ary operator.

Homomorphisms, on the other hand, are defined as follows:

D 2 (Homomorphism). Let $\mathcal{E} = \langle A, F \rangle$ and $\mathcal{B} = \langle B, G \rangle$ be algebras. A mapping $h: \mathcal{E} \rightarrow \mathcal{B}$ is called a *homomorphism* if there is a mapping $h': F \rightarrow G$ such that for all $f \in F$ and all $a_1, \dots, a_n \in A$ holds $h(f(a_1, \dots, a_n)) = h'(f)(h(a_1), \dots, h(a_n))$.

Given this definition, we first have a formal account of what syntax does. It consists of rules that take certain inputs and deliver certain outputs, the outputs being complex expressions, the inputs being ‘parts’ of these. The rules are then represented as operators on the set of syntactic subexpressions.

If we think of a syntactic algebra as a set of expressions of a language upon which a number of operators (syntactic rules) are defined, requiring that these operators always apply to a fixed number of expression and yield a single expression, and allowing that the operators (syntactic rules) may be undefined for certain expression, we get the following partial algebra $\mathbf{E} = \langle E, (F_\gamma)_{\gamma \in \Gamma} \rangle$. Here, E is the set of complex and simple expressions and every F_γ is a partial operator on E with a fixed arity.

Let us now turn to semantics. A meaning assignment is then defined on such a syntactic algebra as a function m from E to M , the set of possible meanings for the expressions of E . Compositionality is then a property of m .

Given one of our F_γ from $\mathbf{E}$, F_k (an k -ary operator on E), m is compositional with respect to this syntactic rule (or ‘ F_k -compositional’) only if there is a partial function from a k -ary partial operator G_k on M (the set of possible meanings) such that whenever $F_k(e_1, \dots, e_k)$ is defined, $m(F_k(e_1, \dots, e_k)) = G_k(m(e_1), \dots, m(e_k))$. We will say that m is *compositional simpliciter* only if m is F_γ -compositional for each F_γ of $\mathbf{E}$. Whenever it is compositional simpliciter, m induces the semantic algebra $\mathbf{M} = \langle M, (G_\gamma)_{\gamma \in \Gamma} \rangle$ on M , and it is a homomorphism between $\mathbf{E}$ and $\mathbf{M}$.

The formal vacuity claim for compositionality can then be proved from a semantic approach as well as from a syntactic approach. Janssen followed a

syntactic route. He proved that if a language is recursively enumerable (there is an algorithm that can generate it), and $m: E \rightarrow M$ a computable function of the expressions of E into M , then there are algebras $\mathbf{E}$ and $\mathbf{M}$ for E and M such that m is a homomorphism. This approach was *syntactic*, for it assumed that no syntactic structure of the language is given beforehand but that we are virtually free to construct the syntactic algebra of the language as we please.

Van Benthem (1984, 1986) strengthened this result somewhat, taking a semantic approach. He assumed that we start with a (possibly partial) term algebra (generated from the set of lexical items by the syntactic operators), that m is defined for the lexical items and that the semantic operations in M which are to correspond to the syntactic operations are arbitrarily fixed as well. But then it is provable in universal algebra that m can be uniquely extended to a homomorphism from our syntactic term algebra to the corresponding semantic algebra (here only the meanings of the lexical items are fixed in advance on the semantic side, we are free to fix the meanings of the complex expressions).⁶

Zadrozny (1994) arrived at his vacuity result also from a semantic perspective. If we begin with a set S of strings generated from an arbitrary alphabet via concatenation (' $s.t$ ' being the concatenation of strings s and t) and consider a meaning function m which assigns the members of another arbitrary set M of meanings to the members of S , he could prove that we can always obtain a new meaning function μ such that for all $s, t \in S$: $\mu(s.t) = \mu(s)(\mu(t))$ and $\mu(s)(s) = m(s)$. This time the assumption was that we are free to choose new meanings (but restrict the syntax considerably and want to be able to retrieve the 'old' meanings from the new ones).⁷

How could these proofs fail? In what sense could they be unconvincing? I think the vacuity proofs can fail to establish the vacuity of the compositionality principle in two main respects.

First of all, all these claims are existence claims. That is they do not by themselves tell us what the compositional meaning function for a language looks like, but only that there is one. Does this establish the vacuity of compositionality? Maybe not necessarily. Here is an argument by Janssen:

The challenge of compositional semantics is not to prove the existence of such a semantics, but to obtain one. The formal results do not help in this respect because the proofs of the theorems assume that some meaning assignment is already given and then turn it into a compositional one. Compositionality is not vacuous, because we have no recipe to obtain a compositional meaning assignment, and because several

⁶For a discussion see Westerståhl (1998).

⁷For a more detailed discussion see Dever (1999), Westerståhl (1998) and Hodges (2001).

proposals are ruled out by the principle. (Janssen, 1997, p. 457)

Janssen then goes on to argue that because the theorems establish the existence of a compositional semantics, the compositionality principle is nevertheless an empirical claim. It was still trivial that there is a compositional semantics, but the non trivial part was to construct one. Compositionality was thus a methodological principle that guides our choices between alternative proposals:

The challenge of semantics is to design a function that assigns meanings, and the present paper argues that the best method is to do so in a compositional way. Compositionality is not an empirical principle, but a methodological one. (Janssen, 1997, p. 457)

I would like to note three things about Janssen's claim: First of all, there is a recipe for obtaining a compositional semantics. This is the recipe behind Zadrozny's result. As noted by Kazmi and Pelletier (1998) and Gendler Szabó (2004), we could simply map every syntax onto itself and thereby obtain an isomorphism (and hence a homomorphism) from syntax to semantics. In case we want the old meanings to be retrievable from the new ones, as in Zadrozny's construction, we also could follow the recipe by Westerståhl: Suppose A is a partial algebra and $m: A \rightarrow M$. Now we generate a new set of meanings $M' = A \times M$. Then $m': A \rightarrow M'$ defined by

D 3. $m'(a) = \langle a, m(a) \rangle$

is compositional, as can easily be seen from the fact that m' is a one-one mapping (from expressions to ordered pairs of expressions and their 'old' meaning). In that case we have again an isomorphism between syntax and semantics and thereby, of course, also a homomorphism as is required by the principle of compositionality. Thus the reason Janssen is giving for why compositionality is not absolutely vacuous seems partly misleading.

Second, and more important, the challenge of finding a compositional meaning assignment is not as unrestricted as assumed in these proofs and it seems *therefore* that the principle is not vacuous. As Westerståhl (1998, p. 641) has argued, none of the mathematical claims discussed so far seem to be of much help for the semanticist.⁸ What semanticists usually are confronted with is data about the meaning of lexical items as well as complex items, data concerning the structure of the language, *plus* the connections between some syntactic operations and some semantic operations being *fixed*. Given that, the conditions

⁸This doesn't mean that the mathematical framework introduced would be useless. A number of very interesting results have been achieved with the help of it, and we will come back to it in the next section. See Pagin (2003a), Westerståhl (1999), Hodges (2001), and Hendriks (2001).

for the proofs of Janssen and Zadrozny to be applicable are never satisfied for the problems of compositionality a semanticist is interested in. For these mathematical proofs we are free in choosing either a semantics or a syntax. But in general we have a (partial) structure on the syntax level, we have intuitions about complex and simple meanings and we have conflicts with compositionality. To show that the language under scrutiny (even if it is human language in general) is compositional is not achieved by throwing the semantic or syntactic data over board (as we would if we took the syntax itself as the semantics for a language), nor by some ad hoc fix by which we would do nothing but construct a *new* meaning function that has the counter instances for compositionality built in unmotivated.

We are not looking for any old meaning function if we evaluate the claim that human languages have or don't have a compositional semantics. What we want is that the meaning function obeys some additional constraints.

The vacuity claims can therefore fail in two respects:

(1) They might fail if they are formulated for just *some* syntactic structure of the language or just *some* assignment of meanings. We are interested in compositionality *given* a structure of a language and *given* certain data about the meanings of simple as well as complex expressions of the language and certain fixed relations between them.

(2) They might fail if they do not deliver a recipe how to obtain a satisfying compositional semantics.

We will not go any deeper into the question of whether the principle of compositionality is vacuous. I pointed out two reasons why one should think that the mathematical arguments given so far are unconvincing and refer to Westerstähl (1998) and Kazmi and Pelletier (1998) for the further substantiation of these reasons. Of course, there is a trivial sense in which every meaning function for a language can be made compositional. This is the sense in which we can hold to every theory we like come what may. However, that is not a special formal feature of compositionality, but of theories in general. We can always choose between giving up a theory, giving up the conflicting data, and even giving up logic.⁹ The question that we will discuss in this paper is *whether there is any good (a priori) reason not to abandon compositionality in cases of conflicting evidence*. The mathematical results do obviously not help here.

The third point I would like to note about Janssen's claim concerns the issue of methodology. If you are a classical philosopher of science¹⁰, you might think that methodological principles are a priori. Maybe that is the way Theo

⁹I will take it for granted that Quine and Duhem convincingly argued for this point.

¹⁰A 'classical philosopher of science' is a philosopher of science who thinks that there is an a priori demarcation between the context of discovery and the context of justification. Methodology is then only concerned with the latter (and a priori), whereas

Janssen thinks about compositionality. In this case we should construct the semantics of the language under consideration compositionally, not because the language under investigation is compositional in some objective sense, but because constructing semantics compositionally serves some other purpose. That the semantics of the language is compositional will then also be *a priori*, given the methodology employed when designing the semantics.

Sometimes the sole reason for employing compositionality as a methodological principle seems to be that doing so seems always possible. Given that the compositionality constraint is almost always satisfiable, we should satisfy it. This *by itself* is certainly *not* a reason to employ a method. If it were *trivial* that for each physical theory T there is a physical theory T' with feature F , and if it were nonetheless non-trivial to construct T' from T , there would be no reason at all to construct a theory with feature F on purpose just because of *that* fact. What is needed instead is a demonstration that having feature F is a good thing, in our case, that compositionality serves some other purpose. It might be sufficient for employing compositionality as a methodological principle if obeying it has in the past led to improvements in the semantics (that seems to be the argument in Janssen, 1997). If that is so, but it cannot be demonstrated how these improvements are systematically¹¹ connected to compositionality, this is clearly not an *a priori* argument for compositionality (and then not our concern in this paper).¹² Then, what could such systematic improvements look like?

Sometimes it is claimed that a compositional semantics reduces complexity. This is true in general, but true because compositionality is a form of systematicity and it is systematicity that reduces complexity (we will come back to that in some detail in the next section).

Compositionality has the additional advantage that the semantic formation rules mirror the syntactic formation rules which is also a kind of complexity reduction. However, there seems to be at least a trade off between the advantage that the semantic value of a complex expression can be read off, in some intuitive sense of 'read off', from its syntactic structure, and the fact that making a non-compositional language compositional might lead to an increase in complexity elsewhere, as seems to be the case when the syntactic structure or formation rules are changed to render a language compositional or when the semantic evaluation is making detours when construed as a function of the meaning of

all contingent, *a posteriori* facts concerning scientific progress belong to the context of discovery.

¹¹Note, however, that it is not necessary for a good methodological rule to be justified that there is an obvious systematic connection with whatever positive thing the employment of the rule causes. The relevant question is whether the rule causes this reliably.

¹²Whether Janssen succeeded to show at all that turning a non-compositional proposal into a compositional one did improve the proposal is doubted by Hodges (1998).

the parts of an expression and the ways they are combined (this seems to be the point of the Sandu/Hintikka response (2001, p. 60) to Hodges compositional semantics for *IF*-languages in which sets of sets of sequences replace mere sets of sequences of the non-compositional semantics (the latter construction, moreover, parallels the semantics of ordinary first-order languages)).

The question is whether there is any much better a priori reason to offer for obeying compositionality as a methodological principle. Here is what can be found in Janssen:

The most valuable arguments [for the compositionality of semantics defined in algebraic terms] are, in my opinion, those concerning the elegance and power of the framework, its heuristic value, and the lack of a mathematically well defined alternative. (Janssen, 1986, p. 38)

The heuristic value is something that sounds interesting. In his earlier publication (1986) Janssen considers only the heuristic benefits of the algebraic framework, but later he generalized this to a claim about the heuristic benefits of compositionality as such. This is a quote from Janssen:

Compositionality is not a formal restriction on what can be achieved, but a methodology on how to proceed. [...] It helps to find weak spots in non-compositional proposals; such proposals have a risk of being defective. Cases where an initially non-compositional proposal was turned into a compositional one, the analysis improved considerably. (Janssen, 1997, p. 461)

Thus far, of course, no argument is given. Every proposal has ‘a risk of being defective’ and the claim about the improvement of analyses is uninteresting as long as there is no systematic connection between compositionality and improvements of a certain kind. Janssen thinks that the connection lies in the fact that compositionality forces you to think about your basic semantic units. Again Janssen:

Compositionality requires a decision on what in a given approach the basic semantic units are: if one has to build meanings from them, it has to be decided what these units are. Compositionality also requires a decision on what the basic units in syntax are and how they are combined. If a proposal is not compositional, it is an indication that the fundamental question what the basic units are, is not answered satisfactorily. If such an answer is provided, the situation under discussion is better understood. So the main reason to follow this methodology, is that compositionality guides research in the right direction. (Janssen, 1997, p. 461)

I am not convinced by this argument. If the methodological guideline was a non-trivial anti-compositionality principle (stated as an imperative) it would equally force the semanticist to make decision about the basic units of syntax and semantics. It seems that even the advice

PoC 3. Think about your basic syntactic and semantic units and the way they are combined, respectively.

is absolutely sufficient to serve the same purpose.

Another argument for the methodological character of the compositionality of meaning departs from a conceptual analysis of meaning. Compositionality might be constitutive of meaning, given that, intuitively, meanings are singled out via the principle of compositionality (in the spirit of David Lewis, 1970: ‘In order to say what a meaning is, first ask what a meaning does, and then find something that does that.’).

Some, then, seem to be convinced that whatever meanings do, they do it compositionally which is an intuition that was maybe even shared by Frege: *Bedeutungen* alone cannot play the role of meanings, for they happen to violate compositionality in intensional contexts. Thus we add *Sinne* to accompany *Bedeutungen* in our semantics and thereby save compositionality. This is – roughly – a story one might tell to reconstruct Frege’s theoretical choices.

If that is true, there is no empirical question of whether or not all possible human languages are compositional. They could not fail to be. In fact, no language natural or artificial could fail to be compositional, because compositionality is a claim about the relation of syntax and semantics, but if compositionality is constitutive for semantics by being a necessary condition for proper meaning functions, there can’t be a non-compositional meaning function.

Does meaning imply compositionality? Hintikka and Sandu (2001) would say it does not, as would Pelletier (1994). Note that everyone who ever denied that a natural language or in fact any language (natural or technical) was compositional fell prey to a conceptual confusion by uttering what is in fact a contradiction in terms. I do not think that this is the correct analysis (by principle of charity).¹³

3 Is Compositionality a *Synthetic A Priori* Truth?

If compositionality is not a vacuous principle there is still a way in which it could be considered *a priori*, given what some philosophers hold about *apriority*.

¹³It seems that the majority of linguists and semanticists would agree that non-compositional meaning functions are conceivable, some – as we know – even think that they are actual phenomena. A theory of concepts that allows that only a minority of an otherwise homogeneous group of competent experts possesses a concept properly seems dubious, to say the least.

If you are a Kantian, for example, you might still think that compositionality can be a priori although it is not vacuous, for you could think that it is a *synthetic* a priori truth. Good empiricists might respond that we can deal with this view rather briefly: there are no synthetic a priori truths, thus if compositionality is not vacuous, it is not a priori. I will not argue for anything else. What will be the subject of this part of my paper are ‘quasi-a priori’-considerations for the compositionality of human languages that are reminiscent of transcendental deductions. They are mere *quasi*-a priori arguments for they do not purport to prove that it is a priori that natural language is compositional. What they try to prove is that natural language must be compositional given empirical facts that are so hard to deny that they are *almost as convincing* as a priori arguments.

What is problematic about transcendental arguments is the status of their premises. Since we are neither interested in Kant nor in the general question of whether transcendental arguments should be regimented that rather than another way, I will take the characterization of transcendental arguments by Roderick Chisholm, show how the standard quasi-a priori arguments for compositionality can be reconstructed as transcendental arguments so characterized, and criticize all of them as unconvincing. I will then try to show that if compositionality is understood as a thesis that quantifies over all possible human languages, it becomes clearer how reasons for or against compositionality could be established by substantial empirical investigation (rather than quasi-a priori arguments).

Transcendental arguments characterized

According to Roderick Chisholm (1978) there are three central features of *transcendental procedures*, the results of which are reported in transcendental arguments. A brief characterization is given in figure 2.

Whether this characterization meets Kant’s own standards for transcendental deductions seems dubious. Kant would probably not have considered every necessary condition for every subject matter in such an argument. He would rather have considered only subject matters which are given to us in special ways and the necessary conditions considered were necessary conditions for our *knowing* of the subject matter in this special way. But that is a side issue for our point. We will only be dealing with arguments that satisfy the characterization given by Chisholm and the conditions under which we want to say that these arguments are justified.

As you can see from the figure, two types of premises are involved in transcendental arguments. First there is the *preanalytic data*. In cases of alleged transcendental deductions this simply might be an empirical claim and is therefore in need of justification. As we will see in a minute, the source of this ‘knowledge’ might well be relevant for the assessment of a quasi-transcendental



Transcendental Arguments (Chisholm Style)

(1) In carrying out the transcendental procedure, one begins with the contemplation of a certain subject-matter. The transcendental procedure makes use of certain propositions about this subject matter. We will say that these propositions constitute the *preanalytic data* of the transcendental procedure.

Preanalytic Data
(self-evident)

(2) As a result of reflecting upon the preanalytic data, one is able to apprehend certain necessary principles about the conditions under which it is possible for the initial subject-matter to exist (or about the conditions under which it is possible for the preanalytic data to be true). We call the second step of the transcendental procedure the *apprehension of transcendental principles*.

Transcendental Principles
(logically necessary)

(3) The transcendental principles are then conjoined with the propositions constituting the preanalytic data and certain consequences are deduced from the result. Given that the premises of the transcendental argument are justified, the significance of the argument will be a function of the significance of these consequences.

Conclusion
(firm as bed-rock)

Figure 2: Kant's transcendental procedure (according to Chisholm).

argument

The second kind of premise is formed by the transcendental principles that state the logically necessary preconditions for the truth of the preanalytic data. Thus, if P is the conjunction of all statements of the first kind in the argument, the transcendental principles are of the form $P \rightarrow Q$, and express truths of logic, broadly conceived (including analytic truth).

Reconstruction of quasi-transcendental arguments

Consider the following (pretty bad) arguments for the compositionality of natural language:

Argument 1

(A1-1) We understand complex expressions by understanding their parts and the way they are combined.

(A1-2) It is necessary for understanding complex expressions in this way that the language is compositional.

(A1-3) Therefore our language is compositional.

Argument 2

(A2-1) We understand complex expressions.

(A2-2) It is necessary for understanding complex expressions that the language is compositional.

(A2-3) Therefore our language is compositional.

Let us begin with argument 2. This argument is obviously unconvincing. It might easily be dismissed by a holistically motivated reply; for it seems conceivable (and might even be an actual phenomenon) that someone manages to learn the meaning of complex expressions without acquiring an understanding of their parts beforehand. Thus we might throw doubt on the alleged necessity of the second premise, the transcendental principle. Consider this argument by Zoltán Gendler Szabó (2000a, pp. 67–68): Arthur might have learned English up to a certain degree and understands the sentences

S 1. It is raining.

and

S 2. This apple is red.

He might also have noticed that adults use S1 and

S 3. Rain is falling.

interchangeably. Thus Arthur knows that he can use S1 and S3 under the exact same circumstances to make correct assertions (and he knows which they are). That seems pretty sufficient for Arthur to have understood S1 as well as S3 (if ‘understanding’ is knowing the truth conditions). This story does not preclude the possibility that Arthur, nevertheless, does not understand the sentence

S 4. This apple is falling.

But in this case S4 is composed of the elements and by the grammatical rules that also compose S1–S3. If that is so, it seems possible that someone comes to an understanding of (at least some) complex expressions without any detour through the understanding of the parts, thus premise (A2-2) seems doubtful.¹⁴

Let us instead turn to argument 1. This only argues that in fact we do understand sentences by understanding their parts and the way they are combined. Doing this, quite obviously, requires that the language allows it, that the language is compositional.¹⁵

This argument is clearly unconvincing because premise (A1-1) smuggles in what is supposed to be established by the argument: it is part of the preanalytic data that the language is compositional if the preanalytic data states that there is a compositional procedure by which we manage to generate the meaning of complex expressions. The transcendental principles are not doing much work in the argument, because compositionality is already guaranteed by the truth of the first premise.

Of course, if the transcendental procedure is successful, the truth of the preanalytic data will entail the compositionality of natural language (via the transcendental principles), but the preanalytic data should not be outright *identical* with a statement claiming the compositionality of language, nor (more importantly) should our reason for believing the preanalytic data to be true, be based on the conviction that natural language is compositional. For the argument to

¹⁴The example simply falsifies the strong compositionality-assumption (A2-2). One might think that in order to do so, we would need an alternative explanation. This is not so. The flight of the bumble-bee might well falsify parts of aerodynamics simply because it is an instance of an intended application and does not behave the way the theory predicts.

¹⁵One might think that this is only true, because (A1-2) is true: that we process a language in a certain way tells us about the way the language is, only because there is a necessary connection between the two ways. But this is besides the point of the example. Premise (A1-1) is supposed to presuppose the conclusion of the argument in a rather blatant way. In other words, the argument does not try to convince someone who doubts that language really is compositional, even granted that we process it that way, but someone who already believes that the question of whether language is compositional is the question of whether we process it that way.

be convincing, we need an *independent* reason to believe the preanalytic data to be true. That is a lesson from the first argument.

Concerning the transcendental principles we observe that their weak spot is their alleged necessity. If a transcendental argument is supposed to be successful, the compositionality of natural language must really be a precondition for the preanalytic data to be true. That is a lesson from the second argument.

With these two considerations in mind, we can now turn to the quasi-a priori arguments for compositionality. Here is such an argument by Donald Davidson¹⁶:

When we regard the meaning of each sentence as a function of a finite number of features of the sentence, we have an insight not only into what there is to be learned; we also understand how an infinite aptitude can be encompassed by finite accomplishments. For suppose that a language lacks this feature; then no matter how many sentences a would-be speaker learns to produce and understand, there will remain others whose meanings are not given by the rules already mastered. It is natural to say that such a language is *unlearnable*. (Davidson, 1984, pp. 8–9)

The argument that this is true for *us* relies on our limitations, all of which we can easily add to the preanalytic data (that man is mortal, that man has finite storage capacity¹⁷). Obviously, compositionality is established by an alleged transcendental argument. We are finite beings but understand infinitely many sentences. The only way that this is possible is that our language is compositional. This argument is sometimes called *the argument from learnability*.

Argument 3 (The Argument from Learnability)

(A3-1) We are able to master infinitely many sentences with different meanings.

(A3-2) We are finite beings.

(A3-3) A noncompositional language with that many sentences is unlearnable for finite beings.

(A3-4) Our language is compositional.

¹⁶Other such arguments were put forward by Noam Chomsky (1980, pp. 76–78) or Jerry Fodor (1987, pp. 147–153).

¹⁷Davidson was aware of this, of course. The quote continues: ‘This argument depends, of course, on a number of empirical assumptions: for example, that we do not at some point suddenly acquire an ability to intuit the meanings of sentences on no rule at all; that each new item of vocabulary, or new grammatical rule, takes some finite time to be learned; that man is mortal.’ (Davidson 1967, p. 9).

Some have remarked that it is not clear that we really master infinitely many sentences, but only potentially infinitely many. But that premise is maybe not even necessary for the argument to go through.¹⁸ It might be sufficient to establish that we understand far more sentences than we could have learned. To see this, consider the following clearly limited case (taken from Grandy, 1990) for which we only consider two noun phrases flanking a transitive verb, as in

S 5. An iguana frightened a tiger.

If we only consider a number of, say, 200 nouns and 50 transitive verbs, we get two million sentences of that form, without even variations in the articles or tense of the verb. Learning another noun would add 20,000 more sentences. Thus it seems not to be of too much relevance for the argument from learnability that we actually cannot master infinitely many sentences.

So far this seems to be a good argument for the compositionality of natural language.¹⁹ However, if we remember our lesson from argument 1, we might ask ourselves what reason we have to believe that premise one of argument 3 is true. That is, what reasons do we have to believe that our language has an infinite or extremely large finite number of meaningful sentences (with different meanings)?

As Peter Pagin has argued (1999, 2002), whatever justification we have to believe that natural language is ‘very rich’ (has a large finite or infinite number of sentences) will undermine the soundness of the argument:

If we are allowed to assume that natural languages are infinitely rich, then we do have a good argument for compositionality. [...] The problem is that we cannot just make the assumption. The claim that natural languages are infinitely rich is a strong claim about natural languages. It would be question begging to simply assume that it is correct. And it is not something we get directly from observations of natural language speakers. It needs a more theoretical justification. (Pagin, 2002, p. 164)

But what could such a *theoretical* justification look like? It might, of course, be the case that we think that natural language is very or even infinitely rich *because* it is compositionally structured²⁰ – that would be the question begging horn of the dilemma. The other horn would be that we believe on the basis of some other feature of our language that it has infinitely many sentences. But in this case, the fact that the language is very rich can’t anymore support that the language is compositional, for in this case we have a reason to believe that some

¹⁸See also Pagin (2002).

¹⁹... at least for a part of it. Of course the language might still have some non-compositional elements. But this is not at issue here.

²⁰... as we just did in the example taken from Grandy.

other mechanism (other than compositionality) is responsible for the richness of our language (generates infinitely many sentences):

Now clearly, if we could justify the assumption that speakers speak an infinitely rich language without compositional structure, then we would know in advance that the learnability argument is flawed, because then, if speakers learn such a language from each other we know that it cannot be correctly explained by means of compositionality. On the other hand, if we can justify directly the assumption that speakers do speak infinitely rich compositional languages, then we already have an argument for compositionality, and need not add any extra consideration about learnability. (Pagin, 2002, p. 164)

Very similar considerations apply to a second standard argument, the argument from new sentences (or ‘the argument from understanding’). This argument one finds often attributed to Frege (for example in Pagin, 2002, p. 166):

It is astonishing what language can do. With a few syllables it can express an incalculable number of thoughts, so that even a thought grasped by a terrestrial being for the very first time can be put into a form of words which will be understood by someone to whom the thought is entirely new. This would be impossible, were we not able to distinguish parts in the thought corresponding to the parts of the sentence, so that the structure of the sentence serves as an image of the structure of the thought. (Frege, 1977, p. 55)

This argument by Frege can be reconstructed in two different ways. If we reconstruct it as an argument from new sentences, it will be possible to give the same reply as we did in response to the argument from learnability:

Argument 4 (The Argument from New Sentences)

(A4-1) Our language has very many sentences with a predetermined meaning.

(A4-2) When we encounter a new sentence of our language that we have never encountered before, we are nonetheless able to understand it.

(A4-3) If our language were non-compositional that could not be explained.

(A4-4) Our language is compositional.

Again it could be asked on what basis we believe in the premise (A4-1). Why should we think that our language has very many sentences with a predetermined meaning? If our reason for believing this is that our language is compositional, why do we need the argument? If, on the other hand, we know of a

mechanism that determines the meaning of very many sentences without being compositional, the argument is flawed.

But, as we've said already, this is only one way to understand Frege's little argument given above. An alternative interpretation is *The Argument from Communication* which is championed by Peter Pagin (1999, 2002, 2003a). The argument is basically an inference to the best explanation (or to the 'only reasonable' explanation).²¹ The preanalytic data it departs from is the observation that communication very often succeeds.

Assume that I happen to be in the city centre of Düsseldorf and for some reason or other do not feel like visiting shoe stores for the rest of the afternoon. To that effect I want to ask my girlfriend (who feels like visiting shoe stores for the rest of the afternoon) to meet me at the town hall half an hour after the shops are closed. I utter the German sentence

S 6. Wir treffen uns 'ne halbe Stunde nach Geschäftsschluss am Rathaus.

Half an hour after the shops are closed on the same day I am standing in front of the town hall, incidentally, my girlfriend happens to be there too. A good explanation for this coincidence is that my girlfriend grasped the thought that I wanted to express with S6. We can assume that she never had heard exactly the same sentence before and that I have never expressed anything with S6 until then, we are both without any previous experience with S6 and no one has explained to us that uttering S6 would be a good way to secure meeting each other in front of a town hall half an hour after the shops are closed.

How come then, that by using [S6] I managed to convert my thought? Could it be hints in the context, or charity of interpretation, or empathy? That any of them, or any combination, could provide the solution, save under special circumstances, is wildly implausible. The only workable explanation is the compositionality explanation, or, more cautiously, an explanation which involves compositionality. (Pagin, 2002, p. 167)

This time the reconstruction of the argument would look somewhat like this:

Argument 5 (The Argument from Communication)

(A5-1) Very often, speakers manage to communicate thoughts by way of uttering sentences that the other party in the communication situation had never heard before.

(A5-2) The only feasible explanation for the success of communication in very many of these cases is that the language is compositional.

²¹In more recent versions of his argument Pagin makes a weaker claim. He now argues that a compositional semantics offers (in relevant cases) a less complex interpretation method and that it is therefore preferable.

(A5-3) Therefore the language is compositional.

The preanalytic data, stated in (A5-1) is not a premise we would need a theoretical reason for. (A5-1) is an empirical fact that we can (and do) have observed. It does not involve any kind of extrapolation or a tacitly built in assumption of compositionality. At least with respect to the first premise, this argument looks much better than the ones we considered so far.

However, this time the trouble is with the transcendental principle, A5-2. The problem we shall discuss now does, of course, also obtain with A3-3 and A4-3 (thus, even if the Pagin response to arguments 3 and 4 did not convince you, the following might). All these premises assume that the only possible way to learn, produce, or understand very (or infinitely) many sentences with different and novel meanings is by way of a compositional semantics.

Again this is very clearly not true. Consider an argument championed by Markus Werning (2004). If we add a rule for holophrastic quotation to a compositional fragment of English and assume that this fragment of English has synonymous expressions, the language will remain productive, of course, but cease to be compositional, if the semantics for holophrastic quotation assign the expression itself as the semantic value of a quoted expression. Hereby, we simply add a syntactic rule for quotation that puts quotation marks around every expression of the language.

D 4. $q: T \rightarrow T, s \mapsto 's'$

The semantic evaluation is then rather simple, we only have to take the expression as the meaning of the expression *with* quotation marks.

D 5. $\mu(q(s)) = s$

Let us assume that 'Lou and Lee are brothers' is synonymous with 'Lee and Lou are brothers' in our compositional fragment of English. Clearly, the extended fragment will be productive, simply because quotation can be iterated (thus it even adds to the productivity), and the meaning function obviously is computable. For the proof of non-compositionality of the extended fragment

we will use ‘ $\ulcorner$ ’ and ‘ $\urcorner$ ’ for meta-linguistic quotation:

$$\mu(\ulcorner \text{Lou and Lee are brothers.} \urcorner) = \mu(\ulcorner \text{Lee and Lou are brothers.} \urcorner) [\text{ass.}] \quad (1)$$

$$\text{Lou and Lee are brothers.} \neq \text{Lee and Lou are brothers.} [\text{ass.}] \quad (2)$$

$$\begin{aligned} \mu(\ulcorner \text{‘Lou and Lee are brothers.’} \urcorner) &= \mu(q(\ulcorner \text{Lou and Lee are brothers.} \urcorner)) \\ &= \ulcorner \text{‘Lou and Lee are brothers.’} \urcorner [\text{D5}] \end{aligned} \quad (3)$$

$$\begin{aligned} \mu(\ulcorner \text{‘Lee and Lou are brothers.’} \urcorner) &= \mu(q(\ulcorner \text{Lee and Lou are brothers.} \urcorner)) \\ &= \ulcorner \text{‘Lee and Lou are brothers.’} \urcorner [\text{D5}] \end{aligned} \quad (4)$$

$$\mu(q(\ulcorner \text{Lou and Lee are brothers.} \urcorner)) \neq \mu(q(\ulcorner \text{Lee and Lou are brothers.} \urcorner)) [2, 3, 4] \quad (5)$$

If we now assume, that the language in question is compositional, there clearly should be a semantic counterpart function μ_q for the syntactic operation q :

$$\mu(q(\ulcorner \text{Lou and Lee are brothers.} \urcorner)) = \mu_q(\mu(\ulcorner \text{Lou and Lee are brothers.} \urcorner)) [\text{comp.}] \quad (6)$$

But then, substitutivity of identicals and another application of compositionality directly leads to an inconsistency:

$$\mu(q(\ulcorner \text{Lou and Lee are brothers.} \urcorner)) = \mu_q(\mu(\ulcorner \text{Lee and Lou are brothers.} \urcorner)) [\text{subst., 1}] \quad (7)$$

$$\mu(q(\ulcorner \text{Lou and Lee are brothers.} \urcorner)) = \mu(q(\ulcorner \text{Lee and Lou are brothers.} \urcorner)) [\text{comp.}] \quad (8)$$

$$\perp \quad [5, 8] \quad (9)$$

Therefore, the extended fragment of English is not compositional.²² But of

²²Of course, there are other ways to analyse quotation, some of which are indeed

course, this language is learnable in some way (you just did), certainly productive, and obviously not compositional.

Explaining mutual understanding of novel sentences

As a more detailed counterexample to the necessity of *compositionality* for the explanation of communication, understanding, and productivity consider Stephen Schiffer's 'Harvey' (1987, pp. 192–207). Schiffer designed his example to show that an explanation of mutual understanding does not even involve reference to semantics. As we will see, this isn't quite correct. However, Schiffer's claim that such an explanation does not need to refer to compositionality can be vindicated.

Harvey is an information processor whose ability to understand novel sentences uttered in (a fragment²³ of) German can be explained without the further assumption that German is compositional. Harvey's beliefs are represented in his belief-box *B* in a lingua mentis *M* (we will take reverse English to represent *M* for the moment) of which, again, nothing is assumed with respect to compositionality. For every possible belief of Harvey there is exactly one sentence *s* of *M* such that Harvey has the belief iff *s* is in *B*. What belief is in *B* is determined by a belief forming mechanism which takes sensory inputs and the present contents of *B* to produce new beliefs in *B*. The unique set of sensory inputs and contents of *B* sufficient for an inner sentence of *M* to be in *B* is that sentence's *conceptual role*.

Given this construction, the question of whether Harvey's understanding of novel German utterances can be explained without recourse to compositionality is then a question of whether the belief forming mechanism can operate given the sensory input of an utterance in German without relying on the compositionality of German at any step. In more detail, the question is whether Harvey can arrive from a sentence of *M*, embedding a representation of the sounds of the utterance, like

S 7. Leinad derettu 'Wir treffen uns'ne halbe Stunde nach Geschaefitsschluss am Rathaus.'

via

S 8. Leinad dias taht ew teem flah na ruoh retfa eht spohs era desolc ta eht nwot llah.

at

compositional. See also Markus Werning's paper in this volume.

²³We will here only consider German without indexicals and without ambiguities. For the extended argument with a fragment of English including these features, see Schiffer (1987, pp. 200–205).

Stephen Schiffer's *Harvey* (inside)

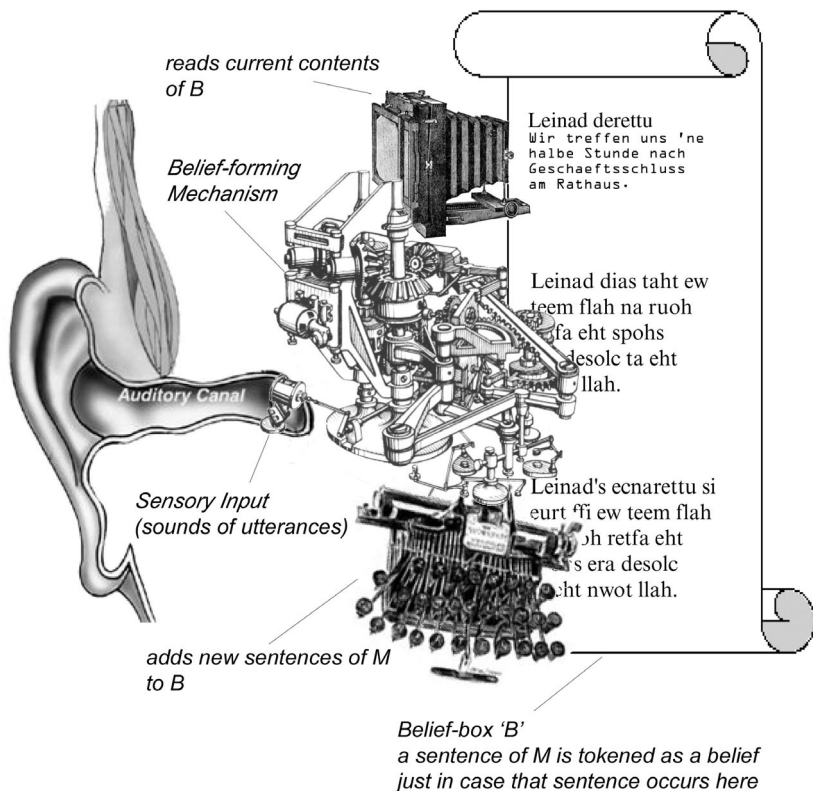


Figure 3: Schiffer's 'Harvey' while processing the sentence I uttered.

S 9. Leinad's ecnarettu si eurt ffi ew teem flah na ruoh retfa eht spohs era desolc ta eht nwot llah.

To make this work, it has to be shown that the conceptual roles of 'dias taht' and 'eurt' do not necessarily presuppose a compositional semantics of either German or *M*. Schiffer shows that this is the case, if there is a recursive function *f* from structural descriptions of sound sequences that are well-formed formulae of German to sentences of *M*, such that (i) *f* is definable in terms of formal features of the expressions in its domain and range, without reference to any semantic features of any expressions in either *M* or German; and (ii) if the referent of the structural description, δ , can be used to say that *p*, then *f*(δ) would token the belief that *p*.

It seems there can well be functions that operate purely syntactical and need not assume the compositionality of *M* or German. In fact, it seems that if *M* represents the structural descriptions of well-formed formulae of German as quoted German sentences *and* Harvey's language of thought *were* German (instead of reverse English), the relevant conceptual roles were trivial and *M* clearly could be non-compositional. Again using '⌈' and '⌋' for meta-linguistic quotation, we can stipulate the conceptual roles for 'hat gesagt' und 'wahr' easily (where 'σ' ranges over sentences of inner and outer (fragmentary) German, and 'Σ' is a structural description of the sentence σ obtained by (some sort of) quotation):

D 6 (Conceptual role of 'hat gesagt'). If the sentence $\lceil \alpha \text{ äußerte } \Sigma \rceil$ is in Harvey's *B*-box, then ceteris paribus, so is $\lceil \alpha \text{ hat gesagt, dass } \sigma \rceil$

D 7 (Conceptual role of 'wahr'). If the sentence $\lceil \alpha \text{ hat gesagt, dass } \sigma \rceil$ is in Harvey's *B*-box, then so is $\lceil \text{Was } \alpha \text{ gesagt hat, (nämlich dass } \sigma) \text{ und damit } \alpha \text{ Äußerung ist wahr gdw. } \sigma \rceil$.

Is this a sufficient explanation of the explanandum? That depends on what we take the explanandum to be. If mutual understanding is what is to be explained, the explanation is not yet satisfying. It does explain why Harvey arrives at *some* interpretation of the uttered sentence, but it does not yet explain why he arrives at the *correct* interpretation, let alone why he *reliably* arrives at the correct interpretation.

As Pagin (2003b) has emphasized, when we seek for an explanation of mutual understanding, we are interested in the latter, an explanation why the hearer arrives at the correct interpretation and why he does so reliably. But this is not explained by Schiffer, who so far only explained why Harvey arrives at a certain sentence in his *B*-box, given a certain input. This is still compatible with complete miscommunication or a mere accidentally correct interpretation of the utterance.

Note that it wouldn't even be sufficient if we had an explanation for successful communication in terms of (a) a recursive procedure by which the hearer

comes to his belief about the content of the utterance, (b) a recursive procedure by which the utterance was composed in accordance with a belief of the speaker, and (c) some connecting fact between the interpretation of the hearer and the intention of the speaker for this utterance token. What we are really looking for, when we ask for an explanation of successful communication, is a deductive-nomological or rather deductive-statistical explanation of a lawlike regularity, viz. that utterance tokens get normally interpreted correctly.²⁴ To provide such an explanation makes necessary to complete Schiffer's explanation with what Pagin calls a 'Content theory', a theory that assigns content to neural sentences on the basis of their syntactic properties in the neural language of interpreter and speaker, respectively.

The Content theory would have to be *systematic* in the following sense: it contains one or more *general* lawlike statements that relate the syntactic properties of neural sentences with their semantic properties. What the Content theory assigns as semantic properties, like truth conditions, to a particular neural sentence then follows from those general lawlike statements together with statements about the particular syntactic properties of the sentence, in conformity with the deductive-nomological or deductive-statistical models of scientific explanation. (Pagin 2003b, 44)

Given such a Content theory, we could explain why a certain thought of the speaker was represented with certain syntactic properties in his neural language, how this led to an utterance of a sentence in the public language, how hearing it caused Harvey to have a neural representation of this sentence with certain syntactic properties and how this sentence of his neural language, eventually, represented a thought that matched the thought of the speakers such that it counts as a correct interpretation of the utterance. A theory that allows for explanations like these is a theory about the meaning of the public language that speaker and hearer use to communicate their thoughts. Part of the theory would explain how syntactic properties of public language sentences connect with thought contents of speakers and interpreters. The lawlikeness of the theory together with the contingent fact that speaker and interpreter are finite beings, assures that such a theory would induce a *systematic* assignment of meanings, determined by the syntactic structure of public language sentences. However, this way we can *only* establish that such a theory must induce a systematic meaning assignment, if there will be a *complete* explanation for communicative success. This does not establish that this meaning assignment must be *compositional*, as we will

²⁴For the type of explanation involved, see Cohnitz (2002). Note that the explanations asked for are explanations of the robustness of a certain phenomenon, rather than a mere explanation of a phenomenon token.

show in the next paragraph, nor does this argument establish that there is such a theory. It might well be that we are interested in the explanation of (apparent) phenomena, although there simply is no (complete) explanation available.²⁵

If the difference between mere systematicity and compositionality is well defined, it seems that the transcendental arguments that rest on quasi-a priori considerations are all²⁶ doomed to fail. What they might establish is merely some kind of ‘grounded recursiveness’ (as Pelletier, 1994 would call it²⁷) or *systematicity* (as defined below). It is sufficient for explaining novelty or productivity if there is some recursive procedure by which we get to the meaning of new expressions. A recursive procedure does not imply that the language be compositional, it will be enough if it is systematic.

Compositionality and mere systematicity

Consider the following example (which is borrowed from Peter Pagin): Suppose we have a language L_1 that consists of expressions generated from the two atomic expressions α, β and the operator σ . α, β are grammatical terms and so is for any grammatical term t , $\sigma(t)$. On the set of grammatical terms we define

²⁵For a discussion, see Cohnitz (2002). In such cases the ‘phenomenon’ is left unexplained, at least in the way it was described. It might then turn out that we must revise the statement describing the phenomenon (turning an apparent non-statistical law into a statistical law, for example). Sometimes, however, there just might be no explanation, simply because we have been mistaken about the nature or even the existence of the explanandum. In the example here, the trouble might be that there is no complete explanation to be had, because what we described as the Content theory might presuppose that we can reduce the mental to the physical and it might turn out that we cannot do so (Pagin, 2003b). But even in this case, a systematic meaning assignment for the public language could explain how it is at all possible that finite beings often have communicative success (Pagin, 2003b). This would not suffice as a complete explanation, but as a ‘how-possibly’ explanation. For the latter notion see Schurz (1995), Hempel (1965).

²⁶The only argument I would know of that is not discussed here but might also be counted as quasi-a priori, is the ‘argument from systematicity’. The trouble with that argument is that it is quite unclear what it is supposed to prove and what exactly its premises are. For a discussion see Gendler Szabó (2004) and the paper by Markus Werning in this volume.

²⁷The difference between Pelletier’s notion of ‘groundedness’ and our notion of ‘systematicity’ defined below is basically that groundedness allows that (aspects of) the context of an utterance may also be part of the recursive procedure by which the meaning of a complex expression is determined, whereas systematicity is defined without mentioning context. We will not go into the question of exactly what is contributed by the context of an utterance and of exactly how that could be reconciled with our notion of systematicity, but leave it for some other occasion.

an enumeration with α as its initial term and the successor operation S :

D 8 (Successor operation S).

1. $S(\alpha) = \beta$,
2. $S(\beta) = \sigma(\alpha)$,
3. $S(\sigma(t)) = \sigma(S(t))$

On the side of the semantics for L_1 , we have a domain of Meanings M which is inductively defined in the following way: there is one basic concept, **lisa**, as well as two primitive functional concepts, **Mother**(x) and **Father**(x) which are the concepts of mother or – in the second case – father of what x is a concept of. The following inductively defines the *rank of a concept*:

D 9 (Rank of a concept).

1. $\text{rank}(c_i) = 0$ if c_i is **lisa**
2. $\text{rank}(c_i) = 1 + \text{rank}(c_1)$ if c_i is **Father**(c_1) or **Mother**(c_1)

Given the rank of concepts, we can define a total ordering on the set of meanings, M :

D 10 (Ordering of meanings).

1. $c_1 < c_2$ if $\text{rank}(c_1) < \text{rank}(c_2)$
2. $c_1 < c_2$ if $\text{rank}(c_1) = \text{rank}(c_2)$, $c_1 = \mathbf{Mother}(c_3)$ and $c_2 = \mathbf{Father}(c_4)$, for some c_3 and c_4
3. $c_1 < c_2$ if $\text{rank}(c_1) = \text{rank}(c_2)$, $c_1 = X(c_3)$, $c_2 = X(c_4)$ and $c_3 < c_4$, for some c_3 and c_4 , where $X = \mathbf{Father}, \mathbf{Mother}$

On top of that we can now define a successor operator **O**:

D 11 (Successor operation **O).** $\mathbf{O}(c_1) = c_2$ iff $c_1 < c_2$ and there is no $c_i \in M$, $c_1 < c_i < c_2$.

and, finally, a meaning function μ :

D 12 (Meaning function).

1. $\mu(\alpha) = \mathbf{lisa}$,
2. $\mu(\beta) = \mathbf{lisa}$,
3. $\mu(S(t)) = \mathbf{O}(\mu(t))$, for $t \neq \alpha$.

The meaning function for L_1 is obviously well-defined, and $\mu(t)$ can be computed for an arbitrary grammatical term t on the basis of its composition. However, μ is not defined by means of recursion over syntax in the normal sense (t does not need to be an immediate constituent of $S(t)$, e.g. in $S(\beta) = \sigma(\alpha)$), and μ is clearly not compositional. Although $\mu(\alpha) = \mu(\beta)$, $\mu(\sigma(\alpha)) \neq \mu(\sigma(\beta))$, since $\mu(\sigma(\alpha)) = \mu(S(\beta)) = \mathbf{O}(\mu(\beta)) = \mathbf{O}(\text{lisa}) = \mathbf{Mother}(\text{lisa})$ whereas $\mu(\sigma(\beta)) = \mu(S(\sigma(\alpha))) = \mu(S(S(\beta))) = \mathbf{O}(\mu(S(\beta))) = \mathbf{O}(\mathbf{O}(\mu(\beta))) = \mathbf{O}(\mathbf{O}(\text{lisa})) = \mathbf{O}(\mathbf{Mother}(\text{lisa})) = \mathbf{Father}(\text{lisa})$.

μ is a *merely systematic* meaning function. The notions of *systematicity*, *compositional systematicity* and *mere systematicity* (as properties of meaning functions) can be defined in the algebraic framework we used in (2.) above, by introducing a slight modification.²⁸ We will now consider a *language algebra* $\mathbf{S}_L$ of a language as a triple $\langle E, \Sigma, A \rangle_L$, where E is the set of expressions of L , Σ a set of functions, each $\sigma \in \Sigma$ a (usually partial) function $E^k \rightarrow E$ for some k , and A a subset of E such that $\Sigma(A) = E$. $\Sigma(A)$ is the set of expressions generated from A by means of Σ . The term algebra $\mathbf{ST}_L$ is also a triple $\langle T, \Sigma^*, A \rangle_L$, where T is the set of terms of S_L denoting elements of E , Σ^* a set of operators associated with Σ , $\sigma \in \Sigma^*$ a (usually partial) function $T^k \rightarrow T$ for some k . The elements of Σ and Σ^* are mapped one-one by $*$.

The set of terms T is then defined inductively together with the valuation function *val*:

D 13 (Set of terms T).

1. every expression $e \in A$ is in T and is an atomic term; $\text{val}(e) = e$,
2. if $t_1 \dots t_n$ are in T and $\sigma_i \in \Sigma$ is defined for $\text{val}(t_1), \dots, \text{val}(t_n)$, then $\sigma_i^*(t_1 \dots t_n)$ is in T , and $\text{val}(\sigma_i^*(t_1 \dots t_n)) = \sigma_i(\text{val}(t_1), \dots, \text{val}(t_n))$,
3. if $\sigma_i^*(t_1 \dots t_n)$ and $\sigma_j^*(u_1 \dots u_m)$ are both in T and $\sigma_i^*(t_1 \dots t_n) = \sigma_j^*(u_1 \dots u_m)$, then $n = m$, $\sigma_i^* = \sigma_j^*$ and $t_i = u_i$, $1 \leq i < n$,

In the example, μ was computable in principle by a finite being, although the set of expressions of L_1 was infinite. If a meaning function allows for that, we will call it *systematic*. To define this property, we will first have to define systematicity for language and term algebras.

D 14. A language algebra $\mathbf{S}_L$ generates the set E of expressions of L from a subset A of E . $\mathbf{S}_L$ (and its associated term algebra $\mathbf{ST}_L$) will be called *systematic* iff $\mathbf{S}_L$ and A_L are both finite and E_L is infinite.

In the example above, we had an algebra $\mathbf{S}_{L_1} = \langle E, \{S\}, \{\alpha, \beta, \sigma\} \rangle_{L_1}$. Consider the alternative *grammatical* algebra for L_1 , $\mathbf{G}_{L_1} = \langle E, \{\sigma\}, \{\alpha, \beta\} \rangle_{L_1}$.

²⁸The following is also in large part due to Peter Pagin.

Both algebras are systematic in the sense defined in (D14). Grammatical algebras are a special case of such systematic algebras and are constrained by the grammaticality restrictions of the language. Term algebras which correspond to grammatical language algebras we will call *grammatical term algebras* (see Hodges, 2001).

Given these notions, we can define *systematicity*, *compositional systematicity* and *mere systematicity* of meaning functions in the following way:

D 15 (Systematicity). A meaning function μ is *systematic* iff μ is a homomorphism from a systematic term algebra $\mathbf{ST}_L = \langle T, \Sigma^*, A \rangle_L$ into a meaning algebra $\mathbf{M}$ of some kind.

D 16 (Compositional systematicity). A meaning function μ is *compositionally systematic* iff μ is a systematic meaning function from a grammatical term algebra $\mathbf{GT}_L = \langle T, \Sigma^*, A \rangle_L$.

D 17 (Mere systematicity). A meaning function μ is *merely systematic* iff μ is a systematic meaning function but is not compositional.

Compositionality, as a property of meaning functions, requires that the recursive procedure by which utterances are produced and interpreted is specific; that it is such that it always only looks at the semantic value of the syntactic parts of an expression and the way they are combined syntactically and computes the semantic value of the complex expression from them. Mere systematicity, again, would allow, for example, for a recursive procedure that requires looking at other simple or complex expressions instead and computing the meaning of the original expression from the syntactic parts of some other sentence and the way *they* are combined. Therefore compositional systematicity is not established by the argument from communication, whereas systematicity might well to be.

Compositionality a posteriori

If compositionality is a substantial claim that is not a priori, we should have some idea how to find evidence in support of compositionality or against it. Compositionality is a thesis about human languages. It is not a thesis about all languages, not even all languages that have certain formal features. I agree with Gendler Szabó (2000a) that compositionality is intended to hold for all possible human languages, whereas ‘possible’ is not meant as it was above, i.e. ‘possible’ in the sense of ‘conceivable’, but ‘possible’ in the sense that we human beings could have developed that language to communicate, given our expressive needs and the structure of our brains. Consider the following necessary conditions for a possible human language:

D 18 (Possible human language). A *possible human language* must be at least (i) a language suitable for the expression and communication of a wide range of

thoughts, and (ii) a language that can be learned by human beings under normal social conditions as a first language.

For illustration of these necessary conditions, Gendler Szabó explains that the language of traffic signs and the language of pure set theory are not possible human languages. Both languages violate (i) in that you can't say in set theory or the language of traffic signs that you have a headache. On the other hand, a language with only two phonemes and a language in which each expression is at least a hundred phonemes long can't be possible human languages because the former's expressions are too easily confused with one another, whereas the expressions of the latter were too hard to keep in mind, thus both would violate (ii).

What the detailed necessary and sufficient conditions for possible human languages are is otherwise²⁹ open to scientific investigation. We do not know yet what languages are all possible in the sense given. Compositionality might then be understood as quantifying over all possible human languages (for more details, see Gendler Szabó 2000a, 2000b):

PoC 4. For every possible human language L and for every complex expression e in L , the meaning of e in L is determined³⁰ by the meanings of the constituents of e in L and by the structure of e in L .

Clearly, positive support for compositionality depends on whether it is necessary for us to use and learn such a language, that this language has certain formal features.

Argument 6 (Advanced Argument from Communication)

(A6-1) Very often, speakers manage to communicate thoughts by way of uttering sentences that the other party in the communication situation had never heard before.

(A6-2) Given the architecture of the human mind, this can only be explained if the native language of those speakers is compositional.

(A6-3) Natural languages (that can serve as a first language) are compositional.

This time the argument does not rest on alleged quasi-a priori assumptions but on outright empirical claims which are far from being self-evident. This holds

²⁹Natural languages will have evolved gradually, so, presumably, there will have been early stages in which you couldn't say that you have a headache in a language that we might nevertheless wish to count as a 'possible human language'. Which criteria we will eventually use for the demarcation is not our concern in this paper.

³⁰Whereas 'determined' should be understood in accordance with our notion of compositional systematicity.

in particular for (A6-2). I do not know whether it is true or false. What we would have to investigate to assess its truth value is whether a language with holophrastic quotation as defined above is a language that can be learned by us as a first language, and is not a language that we can only learn if parasitic upon our mastery of English.³¹ Compositionality is then a significant but extremely general claim about human beings.

4 Concluding Remarks

If compositionality is a very general empirical claim, the way that Gendler Szabó suggests, we can accommodate the following findings:

- We can explain why compositionality is believed to be a general phenomenon rather than a phenomenon of only single languages, like English or German. It is general because the compositionality principle applies to all possible human languages.
- We can explain why not all languages seem to us to be compositional. Some artificial languages are not compositional, some technical extensions of human languages are not compositional.
- We can explain why no a priori argument suffices to establish the truth of the compositionality principle.

One might think that we cannot explain why the principle of semantic compositionality seemed to us to be true, to some of us even to be conceptually true, if the principle is so substantial and the positive reasons for the principle so far away from our actual epistemic situation. Here is Gendler Szabó:

If compositionality turns out to be true, it will seem even more puzzling: why did we have the inclination to believe in it before the real evidence came in? How is it that, even though we are exceedingly uncertain what meaning is, we are convinced that, whatever it is, the meaning of a complex expression supervenes on the meaning of its parts and of its structure? (Gendler Szabó, 2000a, p. 150)

What is called for is a psychological explanation, but I don't see why it should be hard to find one. Why, for example, shouldn't the fact that we understand sentences by understanding their parts and the way they are combined, and a

³¹That is not a problem of formal semantics, of course. The question of whether natural languages are compositional is then a question of cognitive psychology.

sufficient portion of human egocentrism not be enough to account for that, if compositionality is true?³²

Acknowledgements

I would like to thank Wilfrid Hodges, Theo Janssen, Jim Kilbury, Jeff Pelletier, Christoph Rumpf, Markus Werning, and Dag Westerståhl for very helpful discussions of the topics involved in this paper. Special thanks go to Peter Pagin for his comments on earlier versions and his help with part (3.) of this paper.

References

- Baker, G., & Hacker, P. (1980). *Wittgenstein: Understanding and meaning*. Oxford: Blackwell.
- Baker, G., & Hacker, P. (1984a). *Frege: Logical excavations*. Oxford: OUP.
- Baker, G., & Hacker, P. (1984b). *Language, sense & nonsense*. Oxford: Blackwell.
- Benthem, J. van. (1984). The logic of semantics. In F. Landman & F. Veltman (Eds.), *Varieties of formal semantics* (pp. 55–80). Dordrecht: Foris.
- Benthem, J. van. (1986). *Essays in logical semantics* (Vol. 29). Dordrecht: Reidel.
- Burge, T. (1986). Frege on truth. In L. Haaparanta & J. Hintikka (Eds.), *Frege synthesized* (pp. 97–154). Dordrecht: Kluwer.
- Chisholm, R. (1978). What is a transcendental argument? *Neue Hefte für Philosophie*, 14, 19–22.
- Chomsky, N. (1980). *Rules and representations*. Oxford: Basil Blackwell.
- Cohnitz, D. (2002). Explanations are like salted peanuts. on why you can't cut the route towards further reduction. In A. Nimtz, Christian; Beckermann (Ed.), *Argument und Analyse: Proceedings of GAP4*. (pp. 22–36). Mentis.
- Currie, G. (1982). *Frege: An introduction to his philosophy*. Totowa: Barnes & Noble.

³²I'm not saying that there is another argument to the effect that the best explanation for our intuition that The Principle is true is its truth. I'm only saying that if it is true then that might be a good starting point for explaining our intuition, and *that* explanation might then be the best explanation for our intuition *because* the explanans is true.

- Davidson, D. (1984). *Inquiries into truth and interpretation*. Oxford: Clarendon Press.
- Dever, J. (1999). Compositionality as methodology. *Linguistics and Philosophy*, 22, 311–326.
- Dummett, M. (1981a). *Frege: Philosophy of language*. Cambridge (Mass.): Harvard UP.
- Dummett, M. (1981b). *The interpretation of Frege's philosophy*. London: Duckworth.
- Fodor, J. (1987). *Psychosemantics*. Cambridge (Mass.): MIT.
- Fodor, J. (1998). *Concepts: Where cognitive science went wrong*. Oxford: Clarendon Press.
- Fodor, J., & Lepore, E. (2002). *The compositionality papers*. Oxford: Clarendon Press.
- Frege, G. (1977). *Logical investigations*. Oxford: Blackwell.
- Gendler Szabó, Z. (2000a). Compositionality as supervenience. *Linguistics and Philosophy*, 23, 475–505.
- Gendler Szabó, Z. (2000b). *Problems of compositionality*. New York: Garland.
- Gendler Szabó, Z. (Fall 2004). Compositionality. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy*. Stanford: The Metaphysics Research Lab, CSLI Stanford.
- Grandy, R. E. (1990). Understanding and the principle of compositionality. *Philosophical Perspectives*, 4, 557–572.
- Haaparanta, L. (1985a). Frege's context principle. *Communication and Cognition*, 18, 81–94.
- Haaparanta, L. (1985b). Frege's doctrine of being. *Acta Philosophica Fennica*, 39.
- Hempel, C. G. (1965). *Aspects of scientific explanation*. London: The Free Press.
- Hendriks, H. (2001). Compositionality and model-theoretic interpretation. *Journal of Logic, Language, and Information*, 10, 29–84.
- Hodges, W. (1998). Compositionality is not the problem. *Logic and Logical Philosophy*, 6, 7–33.

- Hodges, W. (2001). Formal features of compositionality. *Journal of Logic, Language, and Information*, 10, 7–28.
- Janssen, T. M. V. (1986). *Foundations and applications of Montague grammar, part 1: Philosophy, framework, computer science* (Vol. 1). Amsterdam: CWI.
- Janssen, T. M. V. (1997). Compositionality. In A. t. M. J. van Benthem (Ed.), (pp. 417–473). Amsterdam: Elsevier.
- Janssen, T. M. V. (2001). Frege, contextuality and compositionality. *Journal of Logic, Language and Information*, 10, 115–136.
- Kazmi, A., & Pelletier, F. J. (1998). Is compositionality formally vacuous? *Linguistics and Philosophy*, 21, 629–633.
- Kluge, E.-H. (1980). *The metaphysics of Gottlob Frege*. Nijhoff.
- Machery, E. (2001). Compositionnalite, composition conceptuelle et combinaison prototypique. In *Cognito, Revue francophone internationale en sciences cognitives*, 22, 5–26.
- Montague, R. (1970). Universal grammar. *Theorie*, 36, 373–398.
- Pagin, P. (1999). Radical interpretation and compositional structure. In U. Zeglen (Ed.), (pp. 59–74). London: Routledge.
- Pagin, P. (2002). Rule-following, compositionality and the normativity of meaning. In D. Prawitz (Ed.), *Meaning and interpretation* (Vol. 55, pp. 153–181). Stockholm: Kungl. Vitterhets Historie och Antikvitetsakademien.
- Pagin, P. (2003a). Communication and strong compositionality. *Journal of Philosophical Logic*, 32, 287–322.
- Pagin, P. (2003b). Schiffer on communication. *Facta Philosophica*, 5, 25–48.
- Partee, B. (2004). *Compositionality in formal semantics*. London: Blackwell.
- Pelletier, F. J. (1994). The principle of semantic compositionality. *Topoi*, 13, 11–24.
- Pelletier, F. J. (2001). Did Frege believe Frege’s principle? *Journal of Logic, Language, and Information*, 10, 87–114.
- Resnik, M. (1967). The context principle in Frege’s philosophy. *Philosophy and Phenomenological Research*, 27, 356–365.
- Resnik, M. (1976). Frege’s context principle revisited. In M. Schirn (Ed.), *Studien zu frege iii: Logik und semantik* (pp. 35–49). Stuttgart: Frommann-Holzboog.

- Sandu, G., & Hintikka, J. (2001). Aspects of compositionality. *Journal of Logic, Language, and Information*, 10, 49–61.
- Schiffer, S. (1987). *Remnants of meaning*. Cambridge (Mass.): MIT.
- Schurz, G. (1995). *Wissenschaftliche Erklärung. Ansätze zu einer logisch-pragmatischen Wissenschaftstheorie*. Graz: dbv-Verlag.
- Shwayder, D. (1976). On the determination of reference by sense. In M. Schirn (Ed.), *Studien zu Frege iii: Logik und semantik* (pp. 85–95). Stuttgart: Frommann-Holzboog.
- Skorupski, J. (1984). Dummett's Frege. In C. Wright (Ed.), *Frege: Tradition & influence* (pp. 227–243). Oxford: Blackwell.
- Sluga, H. (1977). Frege's alleged realism. *Inquiry*, 20, 227–242.
- Sluga, H. (1980). *Gottlob Frege*. London: Kegan Paul.
- Sluga, H. (1987). Frege against the Booleans. *NDJFL*, 28, 80–98.
- Werning, M. (2004). *The compositional brain: A unification of conceptual and neuronal perspectives*. Unpublished doctoral dissertation, Philosophical Faculty, Heinrich-Heine University Düsseldorf, Düsseldorf.
- Westerståhl, D. (1998). On mathematical proofs of the vacuity of compositionality. *Linguistics and Philosophy*, 21, 635–643.
- Westerståhl, D. (1999). Idioms and compositionality. In J. Gerbrandy, M. Marx, M. de Rijke, & Y. Venema (Eds.), *JFAK. essays dedicated to Johan van Benthem on the occasion of his 50th birthday*. Amsterdam: Amsterdam University Press.
- Wright, C. (1983). *Frege's conception of numbers as objects*. Aberdeen: Aberdeen UP.
- Zadrozny, W. (1994). From compositionality to systematic semantics. *Linguistics and Philosophy*, 17, 329–342.