1 Background

Compositionality is a property that a language may have and may lack, namely the property that the meaning of any complex expression is determined by the meanings of its parts and the way they are put together. The language can be natural or formal, but it has to be interpreted. That is, meanings, or more generally, semantic values of some sort must be assigned to linguistic expressions, and compositionality concerns precisely the distribution of these values.

Particular semantic analyses that are in fact compositional were given already in antiquity, but apparently without any corresponding general conception. Notions that approximate the modern concept of compositionality did emerge in medieval times. In the Indian tradition, in the 4th or 5th century CE, Šābara says that

The meaning of a sentence is based on the meaning of the words.

and this is proposed as the right interpretation of a sūtra by Jaimini from sometime 3rd-6th century BCE (cf. Houben 1997, 75-76). The first to propose a general principle of this nature in the Western tradition seems to have been Peter Abelard (Abelard forthcoming, 3.00.8) in the first half of the 12th century, saying that

Just as a sentence materially consists in a noun and a verb, so too the understanding of it is put together from the understandings of its parts.\(^2\)

Abelard’s principle directly concerns only subject-predicate sentences, it concerns the understanding process rather than meaning itself, and he is unspecific about the nature of the putting-together operation. The high scholastic conception is different in all three respects. In early middle 14th century John Buridan (Buridan 1998, 2.3, Soph. 2 Thesis 5, QM 5.14, fol. 23vb) states what has become known as the additive principle:

\(^1\)For instance, in *Sophist*, chapters 24-26, Plato discusses subject-predicate sentences, and suggests (pretty much) that such a sentence is true [false] if the predicate (verb) attributes to what the subject (noun) signifies things that are [are not].

\(^2\)Translation by and information from Peter King in King forthcoming, 8.
The signification of a complex expression is the sum of the signification of its non-logical terms.\textsuperscript{3}

The additive principle, with or without the restriction to non-logical terms, appears to have become standard during the late middle ages.\textsuperscript{4} The medieval theorists apparently did not possess the general concept of a function, and instead proposed a particular function, that of summing (collecting). Mere collecting is inadequate, however, since the sentences \textit{All A’s are B’s} and \textit{All B’s are A’s} have the same parts, hence the same collection of part-meanings and hence by the additive principle have the same meaning.

With the development of mathematics and concern with its foundations came a renewed interest in semantics. Gottlob Frege is generally taken to be the first person to have formulated explicitly the notion of compositionality and to claim that it is an essential feature of human language.\textsuperscript{5} In “Über Sinn und Bedeutung”, 1892, he writes:

\begin{quote}
Let us assume for the time being that the sentence has a reference. If we now replace one word of the sentence by another having the same reference, this can have no bearing upon the reference of the sentence.
\end{quote}

(Frege 1892, p. 62)

This is (a special case of) the substitution version of the idea of semantic values being determined; if you replace parts by others with the same value, the value of the whole doesn’t change. Note that the values here are Bedeutungen (referents), such as truth values (for sentences) and individual objects (for individual-denoting terms).

Both the substitution version and the function version (see below) were explicitly stated by Rudolf Carnap in Carnap 1956, p. 121 (for both extension and intension), and labeled ‘Frege’s Principles Interchangeability’. The term ‘compositional’, used in a similar sense, to characterize meaning and understanding, derives from Jerry Fodor and Jerrold Katz Fodor and Katz 1964, with reference to Chomsky but not to Frege or Carnap.

Today, compositionality is a key notion in linguistics, philosophy of language, logic, and computer science, but there are divergent views about its exact formulation, methodological status, and empirical significance. To begin to clarify some of these views we need a framework for talking about compositionality that is sufficiently general to be independent of particular theories of syntax or semantics and yet allows us to capture the core idea behind compositionality.

\textsuperscript{3}Translation by and information from Peter King in King 2001, 4.

\textsuperscript{4}In 1500, Peter of Ailly refers to the common view that it ‘belongs to the [very] notion of an expression that every expression has parts each one of which, when separated, signifies something of what is signified by the whole.’ (Ailly 1980, 30).

\textsuperscript{5}Some writers have doubted that Frege really expressed, or really believed in, compositionality; e.g. Pelletier 2001 and Janssen 2001.
2 A framework

The function version and the substitution version of compositionality are two sides of the same coin: that the meaning (value) of a compound expression is a function of certain other things (other meanings (values) and a ‘mode of composition’). As we will see presently, the substitution version is slightly more general and versatile. To formulate these versions, two things are needed: a set of structured expressions and a semantics for them.

Structure is readily taken as algebraic structure, so that the set $E$ of linguistic expressions is a domain over which certain operations (syntactic rules) are defined, and moreover $E$ is generated by these operations from a subset $A$ of atoms (primitive expressions, e.g. words). In the literature there are essentially two ways of fleshing out this idea. One, which originates with Montague, takes as primitive the fact that linguistic expressions are grouped into categories or sorts, so that a syntactic rule comes with a specification of the sorts of each argument as well as of the value. This use of a many-sorted algebra as an abstract linguistic framework is described in Janssen 1986 and Hendriks 2001. The other approach, first made precise in Hodges 2001, is one-sorted but uses partial algebras instead, so that rather than requiring the arguments of an operation to be of certain sorts, the operation is simply undefined for unwanted arguments. The partial approach is in a sense more general than the many-sorted one, as well as easier to formulate, and we follow it here.

Thus, let a grammar

$$E = (E, A, \Sigma)$$

be a partial algebra, where $E$ and $A$ are as above and $\Sigma$ is a set that, for each required $n \geq 1$, has a subset of partial functions from $E^n$ to $E$, and is such that $E$ is generated from $A$ via $\Sigma$. To illustrate, the familiar rules

$$\begin{align*}
\text{NP} & \to \text{Det N} \\
\text{S} & \to \text{NP VP}
\end{align*}$$

(NP-rule) (S-rule)

correspond to binary partial functions, say $\alpha, \beta \in \Sigma$, such that, if most, dog, and bark are atoms in $A$, one derives as usual the sentence Most dogs bark in $E$, by first applying $\alpha$ to most and dog, and then applying $\beta$ to the result of that and bark. These functions are necessarily partial; for example, $\beta$ is undefined whenever its second argument is dog.

---

6See Montague 1974a, in particular the paper ‘Universal grammar’ from 1970.

7A many-sorted algebra can in a straightforward way be turned into a one-sorted partial one (but not always vice versa), and under a natural condition the sorts can be recovered in the partial algebra. See Westerståhl 2004 for further details and discussion. Note also that some theorists combine partiality with primitive sorts; for example, Keenan and Stabler 2004 and Kracht 2007.

8Note that for a speaker to have a grasp of an infinite syntax by finite means, rules such as the NP-rule and the S-rule must hold in the sense that it is part of a speaker’s competence e.g. that for any pair of terms $(t, u)$ for which the operator $\alpha$ is defined, $\alpha(t, u)$ is an appropriate second argument to $\beta$.

We can call a grammar $E = (E, A, \Sigma)$ inductive if there is a finite partition $(E_s)_{s \in S}$ of
Both in the partial and in the many-sorted framework it may happen that one and the same expression can be generated in more than way, i.e. the grammar may allow structural ambiguity. Also, it may happen that a semantically relevant element is not represented in the surface expression. So in the most general case, it is not really the expressions in $E$ but rather their derivation histories, or ‘analysis trees’, that should be assigned semantic values. These derivation histories are conveniently represented by the terms in the term algebra corresponding to $E$. The sentence itself, i.e. the value of applying the syntactic functions as above, could be identified with a string of words (sounds, phonemes,...), but its derivation history is represented by the term
\[ t = \beta(\alpha(\text{most}, \text{dog}), \text{bark}) \]
in the term algebra. The term algebra is partial too: the grammatical terms are those where all the functions involved are defined for the respective arguments. So $t$ is grammatical but $\beta(\alpha(\text{most}, \text{dog}), \text{dog})$ is not. Let $GT_E$ be the set of grammatical terms for $E$.

Note that the symbols ‘$\alpha$’, ‘$\beta$’... do a double duty here: they name elements of $\Sigma$, i.e., partial functions from expressions to expressions, and these very names are used in the term algebra. For example, we assumed above that
\[ \alpha(\text{most}, \text{dog}) = \text{most dogs} \]
but this equation only makes sense if $\alpha$ is a function, which applied to two elements of $E$ — in this case, the atoms most and dog — yields as value another element of $E$ — in this case, the string most dogs. However, the term $\alpha(\text{most}, \text{dog})$ doesn’t belong to $E$ but to the term algebra. Sometimes one needs to reflect this distinction in the notation; we shall then use symbols with bars over them as names of those symbols. With that notation, we have
\[ \alpha(\text{most}, \text{dog}) \in E \text{ and } \bar{\pi}(\text{most}, \text{dog}) \in GT_E. \]

Each term in $GT_E$ corresponds to a unique string in $E$. Thus, there is a string value function $V$ from $GT_E$ to $E$. $V$ is the identity on atoms, and for complex term $\bar{\pi}(t_1, \ldots, t_n)$ (using now the above notation) we have
\[ V(\bar{\pi}(t_1, \ldots, t_n)) = \alpha(V(t_1), \ldots, V(t_n)), \]

$E$ such that it holds of each $\alpha \in \Sigma$ that its range is a subset of some $E_s$, and its domain is a cartesian product $E_{s_1} \times \cdots \times E_{s_n}$ of sets in $(E_s)_{s \in S}$. That grammars are inductive in this sense is a natural requirement on syntax, but it is not necessary for the semantics to be compositional.

9More correctly, we should write the string value of $\alpha(\text{most}, \text{dog})$ as $\text{most}^{\omega} \text{dogs}$, where ‘$\omega$’ denotes word space and ‘$\cdot$’ concatenation, but the simplified notation used here is easier to read.

10For consistency we could put bars over atomic expressions in $A$ as well to form names of them, thus writing $\bar{\pi}(\text{most}, \text{dog})$. But refraining from this should cause no confusion, since atoms can be assumed to belong to the term algebra too; i.e. they are terms as well as expressions.
where $\alpha$ is defined for the arguments $V(t_1), \ldots, V(t_n)$ precisely when the term $\bar{\alpha}(t_1, \ldots, t_n)$ is grammatical. To illustrate:

$$V(\beta(\bar{\alpha}(\text{most}, \text{dog}), \text{bark})) = \beta(V(\bar{\alpha}(\text{most}, \text{dog})), V(\text{bark}))$$

$$= \beta(\alpha(V(\text{most}), V(\text{dog})), \text{bark})$$

$$= \beta(\alpha(\text{most}, \text{dog}), \text{bark})$$

$$= \beta(\text{most dogs}, \text{bark})$$

$$= \text{most dogs bark}$$

The second thing needed to talk about compositionality is a semantics for $E$. The semantics is naturally taken to be a function $\mu$ to some set $M$ of semantic values (‘meanings’). It is most simple and straightforward to let (a subset of) $GT_E$ be the domain of $\mu$. That is, $\mu$ maps grammatical terms on meanings. That terms are the arguments to $\mu$ does not mean that the expressions themselves are meaningless, only that an expression has meaning derivatively, relative to a way of constructing it, i.e. to a corresponding grammatical term. Indeed, one often slurs over the difference, writing $\mu(e)$ for an expression $e$ in $E$, when what one really should have written is $\mu(t)$ for some grammatical term $t$ with $V(t) = e$.

There are several reasons why the semantic function $\mu$ should be allowed to be partial, too. For example, it may represent our partial understanding of some language, or our attempts at a semantics for a fragment of a language. Further, even a complete semantics will be partial if one wants to maintain a distinction between meaningfulness (being in the domain of $\mu$) and grammaticality (being derivable by the grammar rules).

No assumption is made about meanings. In the abstract framework, the nature of the meanings does not matter more than what is required to determine the relation of synonymy: define, for $u, t \in E$,

$$u \equiv_\mu t \text{ iff } \mu(u), \mu(t) \text{ are both defined and } \mu(u) = \mu(t).$$

$\equiv_\mu$ is a partial equivalence relation on $E$. All that is relevant for compositionality itself is captured by properties of this equivalence relation.

## 3 Variants and properties

### 3.1 Basic compositionality

We can now easily formulate both the function version and the substitution version of compositionality, given a grammar $E$ and a semantics $\mu$ as above.

---

11In other words, $V$ is a homomorphism from the term algebra to the expression algebra $E$.

12There are alternatives. One is to take disambiguated expressions from $E$: expressions somehow annotated to resolve syntactic ambiguities. Phrase structure markings by means of labeled brackets are of this kind. Another option is to have an extra syntactic level, like LF in the Chomsky school, as the semantic function domain. The choice between such alternatives is largely irrelevant from the point of view of compositionality, as long as the syntactic arguments have the required constituent structure.
Funct($\mu$) For every rule $\alpha \in \Sigma$ there is a meaning operation $r_\alpha$ such that if $\alpha(u_1, \ldots, u_n)$ has meaning, $\mu(\alpha(u_1, \ldots, u_n)) = r_\alpha(\mu(u_1), \ldots, \mu(u_n))$.

A variant is to use, for each $n \geq 1$, an $(n+1)$-ary operation $r_n$ instead and let $\alpha$ itself be its first argument, or again a single functional $r$ such that for each $\alpha$, $r(\alpha) = r_\alpha$. Note that Funct($\mu$) presupposes the Domain Principle (DP): subterms of meaningful terms are also meaningful.

The substitution version of compositionality is given by

Subst($\equiv$) If $s[u_1, \ldots, u_n]$ and $s[t_1, \ldots, t_n]$ are both meaningful terms, and if $u_i \equiv_\mu t_i$ for $1 \leq i \leq n$, then $s[u_1, \ldots, u_n] \equiv_\mu s[t_1, \ldots, t_n]$.

Here the notation $s[u_1, \ldots, u_n]$ indicates that the term $s$ contains—not necessarily immediate—disjoint occurrences of subterms among $u_1, \ldots, u_n$, and $s[t_1, \ldots, t_n]$ results from replacing each $u_i$ by $t_i$.\textsuperscript{13} Subst($\equiv$) does not presuppose DP, and one can easily think of semantics for which DP fails. However, a first observation is:

(1) Under DP, Funct($\mu$) and Subst($\equiv$) are equivalent.\textsuperscript{14}

The requirements of basic compositionality are in some respects not so strong, as can be seen from the following observations:

(2) If $\mu$ gives the same meaning to every expression, then Funct($\mu$) holds.
(3) If $\mu$ gives different meanings to all expressions, then Funct($\mu$) holds.

(2) is of course trivial. For (3), consider Subst($\equiv$) and observe that if no two expressions have the same meaning, then $u_i \equiv_\mu t_i$ entails $u_i = t_i$, so Subst($\equiv$), and therefore Funct($\mu$), hold trivially.

3.2 Recursive semantics

The function version of compositional semantics is given by recursion over syntax, but that does not imply that the meaning operations are defined by recursion over meaning, in which case we have recursive semantics. Standard semantic theories are typically both recursive and compositional, but the two

\textsuperscript{13}Restricted to immediate subterms, Subst($\equiv$) says that $\equiv_\mu$ is a (partial) congruence relation:

\begin{align*}
\text{If } & \alpha(u_1, \ldots, u_n) \text{ and } \alpha(t_1, \ldots, t_n) \text{ are both meaningful and } u_i \equiv_\mu t_i \text{ for } 1 \leq i \leq n, \\
\text{then } & \alpha(u_1, \ldots, u_n) \equiv_\mu \alpha(t_1, \ldots, t_n).
\end{align*}

Under DP, this is equivalent to the unrestricted version.

\textsuperscript{14}That Rule($\mu$) implies Subst($\equiv$) is obvious when Subst($\equiv$) is restricted to immediate subterms, and otherwise proved by induction over the complexity of terms. In the other direction, the operations $r_\alpha$ must be found. For $m_1, \ldots, m_n \in M$, let $r_\alpha(m_1, \ldots, m_n) = \mu(\alpha(u_1, \ldots, u_n))$ if there are terms $u_i$ such that $\mu(u_i) = m_i$, $1 \leq i \leq n$, and $\mu(\alpha(u_1, \ldots, u_n))$ is defined. Otherwise, $r_\alpha(m_1, \ldots, m_n)$ can be undefined (or arbitrary). This is enough, as long as we can be certain that the definition is independent of the choice of the $u_i$, but that is precisely what Subst($\equiv$) says.
Notions are mutually independent. For a semantic function $\mu$ to be given by recursion it must hold that:

$$\text{Rec}(\mu) \quad \text{There is a function } b \text{ and for every } \alpha \in \Sigma \text{ an operation } r_\alpha \text{ such that for every meaningful expression } s,$$

$$\mu(s) = \begin{cases} b(s) & \text{if } s \text{ is atomic} \\ r_\alpha(\mu(u_1), \ldots, \mu(u_n), u_1, \ldots, u_n) & \text{if } s = \alpha(u_1, \ldots, u_n) \end{cases}$$

For $\mu$ to be recursive, the basic function $b$ and the meaning composition operation $r_\alpha$ must themselves be recursive, but this is not required in the function version of compositionality. In the other direction, the presence of the terms $u_1, \ldots, u_n$ themselves as arguments to $r_\alpha$, has the effect that the compositional substitution laws need not hold.\footnote{This can happen e.g. with simple semantics for quotation. Cf. Werning 2005.}

Note that if we drop the recursiveness requirement on $b$ and $r_\alpha$, $\text{Rec}(\mu)$ becomes vacuous. This is because $r_\alpha(m_1, \ldots, m_n, u_1, \ldots, u_n)$ can simply be defined to be $\mu(\alpha(u_1, \ldots, u_n))$ whenever $m_i = \mu(u_i)$ for all $i$ and $\alpha(u_1, \ldots, u_n)$ is meaningful (undefined otherwise). Since inter-substitution of synonymous terms changes at least one argument of $r_\alpha$, no counterexample is possible.

### 3.3 Weaker versions

Basic (first-level) compositionality takes the meaning of a complex term to be determined by the meanings of the immediate sub-terms and the top-level syntactic operation. We get a weaker version—second-level compositionality—if we require only that the operations of the two highest levels, together with the meanings of terms at the second level, determine the meaning of the whole complex term.\footnote{A possible example, from Peters and Westerståhl 2006, ch. 7, concerns possessive determiner phrases like some student’s, taken to be generated by (NP-rule) above and

$$\text{Det} \rightarrow \text{NP} \, \text{’s} \quad \text{(Poss)}$$

If the semantic value of the Det in (NP-rule) is a type (1, 1) quantifier $Q$ and the value of $N$ is a set $C$, the value of the resulting NP is arguably the type (1) quantifier $QC$, i.e. $Q$ with its restriction argument frozen to $C$. Peters and Westerståhl argue that when this complex NP is the argument in the (Poss) rule, a determination of the semantic value of the resulting possessive determiner requires access to both $Q$ and $C$, but they also show that these cannot always be recovered from $QC$. If so, the corresponding semantics is second-level but not (first-level) compositional.}

Third-level compositionality is defined analogously, and is weaker still. In the extreme case we have bottom-level, or weak functional compositionality, if the meaning the complex term is determined only by the meanings of its atomic constituents and the total syntactic construction (i.e. the derived operation that is extracted from a complex term by knocking out the atomic constituents). A function version of this is somewhat cumbersome to formulate precisely (but see Hodges 2001, sect. 5),\footnote{Terminology concerning compositionality is somewhat fluctuating. David Dowty in Dowty 2007 calls (an approximate version of) weak functional compositionality \textit{Frege’s Principle}, and refers to $\text{Funct}(\mu)$ as \textit{homomorphism compositionality}, or \textit{strictly local compositionality}, whereas the substitution version}
becomes simply:

\[\text{AtSubst}(\equiv \mu)\]

Just like \(\text{Subst}(\equiv \mu)\) except that the \(u_i\) and \(t_i\) are all atomic.

Although weak compositionality is not completely trivial (a language could lack the property), it does not serve the language users very well: the meaning operation \(r_\alpha\) that corresponds to a complex syntactic operation \(\alpha\) cannot be predicted from its build-up out of simpler syntactic operations and their corresponding meaning operations. Hence, there will be infinitely many complex syntactic operations whose semantic significance must be learned one by one.

### 3.4 Stronger versions

We get stronger versions of compositionality by enlarging the domain of the semantic function, or by placing additional restrictions on meaningfulness or on meaning composition operations. An example of the first is Zoltan Szabo’s idea (Szabó 2000) that the same meaning operations define semantic functions in all possible human languages, not just for all sentences in each language taken by itself. That is, whenever two languages have the same syntactic operation, they also associate the same meaning operation with it.

An example of the second option is what Wilfrid Hodges has called the *Husserl property* (going back to ideas in Husserl, Husserl 1900):

(\text{Huss}) Synonymous terms belong to the same (semantic) category.

Here the notion of category is defined in terms of substitution; say that \(u \sim_\mu t\) if, for every term \(s\) in \(E\), \(s[u] \in \text{dom}(\mu)\) iff \(s[t] \in \text{dom}(\mu)\). So (Huss) says that synonymous terms can be inter-substituted without loss of meaningfulness. This is often a reasonable requirement. (Huss) also has the consequence that \(\text{Subst}(\equiv \mu)\) can be simplified to \(\text{Subst}_1(\equiv \mu)\), which only deals with replacing one subterm by another. Then one can replace \(n\) subterms by applying \(\text{Subst}_1(\equiv \mu)\) \(n\) times; (Huss) guarantees that all the ‘intermediate’ terms are meaningful.

An example of the third kind is that of requiring the meaning composition operations to be recursive, or computable. To make this idea more precise, in analogy with arithmetic, we need to impose more order on the meaning domain. We have to view meanings as themselves given by an algebra \(M = (M, B, \Omega)\), where \(B \subseteq M\) is a finite set of basic meanings, \(\Omega\) is a finite set of elementary operations from \(n\)-tuples of meanings to meanings, and \(M\) is generated from \(B\) by means of the operations in \(\Omega\). This allows the definition of functions by recursion over \(M\), and the meaning operations are to be of this kind (those in \(\Omega\) will correspond to the successor operation for ordinary recursion over natural numbers). The semantic function \(\mu\) is then defined simultaneously by recursion over syntax and by recursion over the meaning domain. Assuming that the

or context-free semantics. In Larson and Segal 1995, this is called strong compositionality. The labels second-level compositionality, third-level, etc. are not standard in the literature but seem appropriate.
elementary meaning operations are computable in a sense relevant to cognition, the semantic function itself is computable.

A further step in this direction is to require that the meaning operation relevant to semantics are of some restricted kind that makes them easy to compute, and thereby reduces or minimizes the (time) complexity of semantic interpretation. For instance, meaning operations that are either elementary or else formed from elementary operations by function composition and function application would be of this kind.18

Another strengthening, also introduced in Hodges 2001, concerns Frege’s so-called Context Principle. A famous but cryptic saying by Frege in Frege 1884 is: “Never ask for the meaning of a word in isolation, but only in the context of a sentence” (p. x). This principle has been much discussed in the literature19, and often taken to conflict with compositionality. However, if not seen as saying that words somehow mysteriously lose their meaning in isolation, it can be interpreted as a constraint on meanings, in the form of what we might call the Contribution Principle, roughly:

\[(CP) \quad \text{The meaning of a term is the contribution it makes to the meanings of complex terms of which it is a part.}\]

This is still vague, but Hodges notes that it can be made precise in the form of an additional requirement on the synonymy \(\equiv_\mu\). Assuming (Huss), as Hodges does here, consider:

\[\text{InvSubst}_\equiv(\equiv_\mu) \quad \text{If } u \not\equiv_\mu t \text{ then there is some term } s \text{ such that either exactly one of } s[u] \text{ and } s[t] \text{ are meaningful, or both are and } s[u] \not\equiv_\mu s[t].\]

This entails that if two terms of the same category are such that no complex term of which the first is a part changes meaning when the first is replaced by the second, they are synonymous. That is, if they make the same contribution to all such complex terms, their meanings cannot be distinguished. This can be taken as one half of (CP), and compositionality in the form of Subst\(_1(\equiv_\mu)\) as the other.20

We can take a step further in this direction by requiring that substitution of terms by terms with different meanings always changes meaning:
InvSubst$_\forall(\equiv_\mu)$ If for some $i$, $0 \leq i \leq n$, $u_i \not\equiv_\mu t_i$, then for every term $s[u_1, \ldots, u_n]$ it holds that either exactly one of $s[u_1, \ldots, u_n]$ and $s[t_1, \ldots, t_n]$ are meaningful, or both are and $s[u_1, \ldots, u_n] \not\equiv_\mu s[t_1, \ldots, t_n]$.

This principle disallows synonymy between complex terms that can be transformed into each other by substitution of constituents at least some of which are non-synonymous, but it allows two terms with different structure to be synonymous. Carnap’s principle of synonymy as intensional isomorphism forbids this, too. With the concept of intension from possible-worlds semantics it can be stated as

(RC) $t \equiv_\mu u$ iff

(i) $t, u$ are atomic and co-intensional, or

(ii) for some $\alpha, t = \alpha(t_1, \ldots, t_n), u = \alpha(u_1, \ldots, u_n)$, and $t_i \equiv_\mu u_i, 1 \leq i \leq n$

(RC) entails both Subst$(\equiv_\mu)$ and InvSubst$_\forall(\equiv_\mu)$, but is very restrictive. It disallows synonymy between brother and male sibling as well as between John loves Susan and Susan is loved by John, and allows different terms to be synonymous only if they differ at most in being transformed from each other by substitution of synonymous atomic terms.

This seems too strong. We get an intermediate requirement as follows. First we define two terms $t$ and $u$ to be $\mu$-congruent, $t \simeq_\mu u$:

$\simeq_\mu$ $t \simeq_\mu u$ iff

(i) $t$ or $u$ is atomic, $t \equiv_\mu u$, and neither is a constituent of the other, or

(ii) $t = \alpha(t_1, \ldots, t_n), u = \beta(u_1, \ldots, u_n), t_i \simeq_\mu u_i, 1 \leq i \leq n$, and for all $s_1, \ldots, s_n, \alpha(s_1, \ldots, s_n) \equiv_\mu \beta(s_1, \ldots, s_n)$, if either is defined.

Then we require synonymous term to be congruent:

(Cong) If $t \equiv_\mu u$, then $t \simeq_\mu u$.

By (Cong), synonymous terms cannot differ much syntactically, but they may differ in the two crucial respects forbidden by (RC). (Cong) does not hold for natural language if logically equivalent sentences are taken as synonymous, but that it holds otherwise remains an hypothesis.

It is a consequence of (Cong) that meanings are structured entities or can be represented as structured entities, i.e. entities uniquely determined by how they are built, i.e. again entities from which constituents can be extracted. That is, we have projection operations:

(Rev) For every meaning operation $r : E^n \rightarrow E$ there are projection operations $s_{r,i}$ such that $s_{r,i}(r(m_1, \ldots, m_n)) = m_i$.

(Rev) alone tells us nothing about the semantics. Only together with the fact
that the operations $r_i$ are meaning operations for a compositional semantic function $\mu$ do we get semantic consequences. The main consequence is that we also have a kind of inverse functional compositionality:

\[
\text{InvFunct}(\mu)
\]

The syntactic expression of a complex meaning $m$ is determined, up to $\mu$-congruence, by the composition of $m$ and the syntactic expressions of its parts.

For the philosophical significance of inverse compositionality, see sections 4.6 and 5.2 below.\textsuperscript{21}

### 3.5 Direct and indirect compositionality

The terms or derivation trees that are the arguments of the semantic function may differ more or less from the expressions (strings of symbols) that correspond to them. In Jacobson 2002, Pauline Jacobson distinguishes between direct and indirect compositionality, according to the relation between terms and expressions, as well as between strong direct and weak direct compositionality. Informally, in strong direct compositionality, expressions are built up from subexpressions simply by means of concatenation, left or right. In weak direct compositionality, one expression may wrap around another (as \textit{call up} wraps around \textit{him} in \textit{call him up}). As we understand Jacobson, the following defines her strong direct compositionality. Let $V(t)$ (as before) be the expression (string) that corresponds to the grammatical term $t$, and likewise the occurrence of a string that corresponds to the occurrence of a term in a larger term. Distinct occurrences of terms correspond to distinct occurrences of strings. Then we can state:

\[
\text{(SDC)} \quad \text{A language is strongly directly compositional iff}
\]

i) For any subterm occurrence $t'$ of a complex grammatical term $t$, $V(t')$ is a substring occurrence of $V(t)$, and

ii) for every symbol occurrence $x$ in $V(t)$ there is a proper subterm $t''$ of $t$ such that $x$ is in $V(t'')$.

iii) There is a (total) compositional semantic function $\mu$ defined on the grammatical terms.\textsuperscript{22}

The weak direct version is like the strong version except that substrings are allowed to have discontinuous occurrences: every symbol occurrence in the contained string has an occurrence in the containing string and the order between symbol occurrences is preserved, but symbol occurrences from other string occurrences may intervene. For indirect compositionality, i.e. for our notion of

\textsuperscript{21}For $(\simeq_\mu)$, $(\text{Cong})$, $\text{InvFunct}(\mu)$, and a proof that (Rev) is a consequence of (Cong) (really of the equivalent statement that the meaning algebra is a free algebra), see Pagin 2003a. (Rev) seems to be what Jerry Fodor understands by ‘reverse compositionality’ in e.g. Fodor 2000, p. 371.

\textsuperscript{22}Note that for Jacobson, as for Kracht (see next subsection), the arguments of the semantic function are really grammatical terms formed from expression triples (phonology, category, meaning), but this does not essentially change the situation.
compositionality here, both conditions i) and ii) (as well as the totality requirement on \( \mu \)) are dropped: syntactic operations may delete strings, reorder strings, make substitutions and add new elements. In addition, Jacobson distinguishes as more radically indirect theories in which the arguments to the semantic function belong to an indirectly derived syntactic level, like LF in the Chomsky school.

Strictly speaking, the direct/indirect distinction is not a distinction between kinds of compositionality, but between kinds of syntax. Still, discussion of it tends to focus on the role of compositionality in linguistics, e.g. whether to let the choice of syntactic theory be guided by compositionality (cf. Dowty 2007 and Kracht 2007).

3.6 Expression triples

Some linguists, among them Jacobson, tend to think of grammar rules as applying to \textit{signs}, where a sign is a triple \((e, k, m)\) consisting of a string, a syntactic category, and a meaning. This is formalized by Marcus Kracht (see Kracht 2003, Kracht 2007), who defines an \textit{interpreted language} to be a set \(L\) of signs in this sense, and a \textit{grammar} \(G\) as a set of partial functions (of various arities) from signs to signs, such that \(L\) is generated by the functions in \(G\) from a subset of atomic (lexical) signs. Thus, a meaning assignment is built into the language, and grammar rules are taken to apply to meanings as well.

This looks like a potential strengthening of our notion of grammar, but is not really used that way, partly because the grammar is taken to operate independently (though in parallel) at each of the three levels. Let \(p_1, p_2,\) and \(p_3\) be the projection functions on triples yielding their first, second, and third elements, respectively. Kracht calls a grammar \textit{compositional} if for each \(n\)-ary grammar rule \(\alpha\) there are three operations \(r_{\alpha,1}, r_{\alpha,2},\) and \(r_{\alpha,3}\) such that for all signs \(\sigma_1, \ldots, \sigma_n\) for which \(\alpha\) is defined,

\[
\alpha(\sigma_1, \ldots, \sigma_n) = \langle r_{\alpha,1}(p_1(\sigma_1), \ldots, p_1(\sigma_n)), r_{\alpha,2}(p_2(\sigma_1), \ldots, p_2(\sigma_n)), r_{\alpha,3}(p_3(\sigma_1), \ldots, p_3(\sigma_n)) \rangle
\]

and moreover \(\alpha(\sigma_1, \ldots, \sigma_n)\) is defined if and only if each \(r_{\alpha,i}\) is defined for the corresponding projections.

In a sense, however, this is not really a variant of compositionality but rather another way to organize grammars and semantics. This is indicated by (4) and (5) below, which are not hard to verify.\(^{24}\) First, call \(G\) \textit{strict} if \(\alpha(\sigma_1, \ldots, \sigma_n)\)

\(^{23}\)For discussions of the general linguistic significance of the distinction, see Barker and Jacobson 2007.

\(^{24}\)It may seem more natural to extract from \(G\) a semantics in the sense of a function from \textit{strings} to meanings, rather than from signs to meanings as in (4). This, however, cannot be done without extra assumptions on \(G\). For example, one might want \(G\) to allow for \textit{ambiguity}, i.e. the possibility of \(\sigma = (e, k, m)\) and \(\sigma' = (e, k, m')\) belonging to \(L\) while \(m \neq m'\); here \(\sigma\) and \(\sigma'\) may be atomic (lexical ambiguity) or even complex with the same derivation history; cf. section 6.4 below. This would be a use of Kracht’s format going beyond the organization of grammars and semantics used here, and would exclude a functional assignment of meanings to strings.
defined and \( p_1(\tau_i) = p_1(\sigma_i) \) for \( 1 \leq i \leq n \) entails \( \alpha(\tau_1, \ldots, \tau_n) \) defined, and similarly for the other projections. All compositional grammars are strict.

4) Every grammar \( G \) in Kracht’s sense for an interpreted language \( L \) is a grammar \((E, A, \Sigma)\) in the sense of section 2 (with \( E = L \), \( A = \) the set of atomic signs in \( L \), and \( \Sigma = \) the set of partial functions of \( G \)). Provided \( G \) is strict, \( G \) is compositional (in Kracht’s sense) iff each of \( p_1, p_2, \) and \( p_3, \) seen as assignments of values to signs (so \( p_3 \) is the meaning assignment), is compositional (in our sense).

5) Conversely, if \( E = (E, A, \Sigma) \) is a grammar and \( \mu \) a semantics for the grammatical terms of \( E \), let \( L = \{\langle u, u, \mu(u) \rangle : u \in \text{dom}(\mu)\} \). Define a grammar \( G \) for \( L \) (with the obvious atomic signs) by letting

\[
\alpha(\langle u_1, u_1, \mu(u_1) \rangle, \ldots, \langle u_n, u_n, \mu(u_n) \rangle) = \\
\langle \alpha(u_1, \ldots, u_n), \mu(\alpha(u_1, \ldots, u_n)) \rangle
\]

whenever \( \alpha \in \Sigma \) is defined for \( u_1, \ldots, u_n \) and \( \alpha(u_1, \ldots, u_n) \in \text{dom}(\mu) \) (undefined otherwise). Provided \( \mu \) is closed under subterms and has the Husserl property, \( \mu \) is compositional iff \( G \) is compositional.\(^{25}\)

3.7 Context dependence

In standard possible-worlds semantics the role of meanings are served by the intensions, i.e. functions from possible worlds to extensions. For instance, the intension of a sentence \( s \), \( I(s) \) is a function that for a possible world \( w \) as argument returns a truth value, if the function is defined for \( w \). Montague (Montague 1974b) extended this idea to include not just worlds but arbitrary indices \( i \) from some set \( I \), as ordered \( n \)-tuples of contextual factors that are relevant to semantic evaluation. Time and place of utterance are typical elements in such indices. The semantic function \( \mu \) assigns a meaning to an term \( t \), such that \( \mu(t) \) itself is a function \( \mu_t \) such that for an index \( i \in I \) \( \mu_t(i) \) gives an extension as value.

For such an apparatus, the concept of compositionality can be straightforwardly applied. The situation gets more complicated when the semantic function itself takes contextual arguments, e.g. if a meaning-in-context for a term \( t \) in context \( c \) is given as \( \mu(t, c) \). The reason for such a change might be the view that the contextual meanings are contents in their own right, not just extensional fall-outs of the standing, context-independent meaning. But with context as a separate argument to the semantic function, we have a new source of variation. The most natural extension of compositionality to such a context semantics is given by

\[ \text{C-Funct}(\mu) \]

For every rule \( \alpha \in \Sigma \) there is a meaning operation \( r_\alpha \) such that for every context \( c \), if \( \alpha(u_1, \ldots, u_n) \) has meaning in \( c \), then

\[
\mu(\alpha(u_1, \ldots, u_n), c) = r_\alpha(\mu(u_1, c), \ldots, \mu(u_n, c)).
\]

\(^{25}\)Here we have secured strictness of \( G \) by letting each term \( u_i \) make up its own grammatical category. If the grammar \( E \) is inductive in the sense of footnote 8, we can instead more naturally assign categories that correspond to the partition sets.
C-Funct(\(\mu\)) seems like a straightforward extension of compositionality to a contextual semantics, but it can fail in a way non-contextual semantics cannot, by a context-shift failure. For we can suppose that although \(\mu(u_i, c) = \mu(u_i, c')\), \(1 \leq i \leq n\), we still have \(\mu(\alpha(u_1, \ldots, u_n), c) \neq \alpha(u_1, \ldots, u_n, c')\). One could claim that this is a possible result of so-called unarticulated constituents. Maybe the meaning of the sentence

(6) It rains

is sensitive to the location of utterance, while none of the constituents of that sentence (say, it and rains) is sensitive to location. Then the contextual meaning of the sentence at a location \(l\) is different from the contextual meaning of the sentence at another location \(l'\), even though there is no such difference in contextual meaning for any of the parts (cf. Perry 1986). This may hold even if substitution of expressions is compositional.

There is therefore room for a weaker principle that cannot fail in this way, where the meaning operation itself takes a context argument:

\[
\text{C-Funct}(\mu)_c \quad \text{For every rule } \alpha \in \Sigma \text{ there is a meaning operation } r_\alpha \text{ such that for every context } c, \text{ if } \alpha(u_1, \ldots, u_n) \text{ has meaning in } c, \text{ then } \\
\mu(\alpha(u_1, \ldots, u_n), c) = r_\alpha(\mu(u_1, c), \ldots, \mu(u_n, c), c).
\]

The only difference is the last argument of \(r_\alpha\). Because of this argument, C-Funct(\(\mu\)) is not sensitive to the counterexample above, and is more similar to non-contextual compositionality in this respect.

Standing meanings can be derived from contextual meanings by abstracting over the context argument: \(\mu_s(t) = \lambda c(\mu(t, c))\), where \(\mu_s\) is the semantic function for standing meaning. It can be shown that if \(\mu\) obeys C-Funct(\(\mu\)) or C-Funct(\(\mu\)), then \(\mu_s\) obeys Funct(\(\mu\)). That is, compositionality for contextual meaning entails compositionality for standing meaning. The converse does not hold, for we can let \(\mu(t, c) = \mu(u, c)\), while \(\mu(t, c') \neq \mu(u, c')\). Then, if \(\mu(\alpha(t, c) \neq \mu(\alpha(u, c))\), we have substitution failure in context \(c\) for contextual meaning, but \(t\) and \(u\) cannot yield substitution failure for standing meaning, since their standing meanings are different (cf. Pagin 2005).

So far, we have been concerned with extra-linguistic context, but we can also extend compositional semantics to dependence on linguistic context. That is, the semantic value of some particular occurrence of an expression may depend on whether that is an occurrence in, say, an extensional context, or an intensional context, or a hyperintensional context, a quotation context, or yet something else.

A framework for such a semantics needs a set \(C\) of context types, an initial null context type \(\theta \in C\) for unembedded occurrences (i.e. terms simpliciter), and a binary function \(\psi\) from context types and syntactic operators to context types. If a term \(\alpha(t_1, \ldots, t_n)\) occurs in a context type \(c_i\), then the terms \(t_1, \ldots, t_n\) will occur in a context of type \(\psi(c_i, \alpha)\). The context type for a particular occurrence \(t_i^\alpha\) of a term \(t_i\) in a host term \(t\) is then determined by its immediately embedding operator \(\alpha_1\), its immediately embedding operator, and so on until the topmost
operator occurrence.

The semantic function \( \mu \) takes a term \( t \) and a context type \( c \) into a semantic value. The only thing that will differ in the clause for a complex term for linguistic context from C-Funct(\( \mu \)) above is that the context of the subterms may be different, according to the function \( \psi \), from the context of the containing term:

\[
\text{LC-Funct}(\mu)_c \quad \text{For every rule } \alpha \in \Sigma \text{ there is a meaning operation } r_\alpha \text{ such that for every context } c, \text{ if } \alpha(u_1, \ldots, u_n) \text{ has meaning in } c, \text{ then }

\mu(\alpha(u_1, \ldots, u_n), c) = r_\alpha(\mu(u_1, c'), \ldots, \mu(u_n, c'), c), \text{ where } c' = \psi(c, \alpha).
\]

This will be applied to a semantics for quotation contexts in section 6.2 below.

4 Arguments in favor of compositionality

4.1 Learnability

Perhaps the most common argument for compositionality is the argument from learnability: A natural language has infinitely many meaningful sentences. It is impossible for a human speaker to learn the meaning of each sentence one by one. Rather, it must be possible for a speaker to learn the entire language by learning the meaning of a finite number of expressions, and a finite number of construction forms. For this to be possible, the language must have a compositional semantics. The argument was to some extent anticipated already in Sanskrit philosophy of language.\(^{26}\) A modern classical passage plausibly interpreted along these lines is due to Donald Davidson:

It is conceded by most philosophers of language, and recently by some linguists, that a satisfactory theory of meaning must give an account of how the meanings of sentences depend upon the meanings of words. Unless such an account could be supplied for a particular language, it is argued, there would be no explaining the fact that we can learn the language: no explaining the fact that, on mastering a finite vocabulary and a finite set of rules, we are prepared to produce and understand any of a potential infinitude of sentences. I do not dispute these vague claims, in which I sense more than a kernel of truth. Instead I want to ask what it is for a theory to give an account of the kind adumbrated (Davidson 1967, 17).

Properly spelled out, the problem is not that of learning the meaning of infinitely many meaningful sentences (given that one has command of a syntax), for if I learn that they all mean that \textit{snow is white}, I have already accomplished the task. Rather, the problem is that there are infinitely many propositions

\(^{26}\)During the first or second century BC Patañjali writes: ‘...Bṛhaspati addressed Indra during a thousand divine years going over the grammatical expressions by speaking each particular word, and still he did not attain the end. ...But then how are grammatical expressions understood? Some work containing general and particular rules has to be composed ...’ Cf. Staal 1969, 501-02. Thanks to Brendan Gillon for the reference.
that are each expressed by some sentence in the language (with contextual parameters fixed), and hence infinitely many equivalence classes of synonymous sentences.

Still, as an argument for compositionality, the learnability argument has two main weaknesses. First, the premise that there are infinitely many sentences that have a determinate meaning although they have never been used by any speaker, is a very strong premise, in need of justification. That is, at a given time $t_0$, it may be that the speaker or speakers employ a semantic function $\mu$ defined for infinitely many sentences, or it may be that they employ an alternative function $\mu_0$ which agrees with $\mu$ on all sentences that have in fact been used but is simply undefined for all that have not been used. On the alternative hypothesis, when using a new sentence $s$, the speaker or the community gives some meaning to $s$, thereby extending $\mu_0$ to $\mu_1$, and so on. Phenomenologically, of course, the new sentence seemed to the speakers to come already equipped with meaning, but that was just an illusion. On this alternative hypothesis, there is no infinite semantics to be learned. To argue that there is a learnability problem, we must first justify the premise that we employ an infinite semantic function.27 This cannot be justified by induction, for we cannot infer from finding sentences meaningful that they were meaningful before we found them, and exactly that would have to be the induction base.

The second weakness is that even with the infinity premise in place, the conclusion of the argument would be that the semantics must be computable, but computability does not entail compositionality, as we have seen.

4.2 Novelty

Closely related to the learnability argument is the argument from novelty: speakers are able to understand sentences they have never heard before, which is possible only if the language is compositional.

When the argument is interpreted so that, as in the learnability argument, we need to explain how speakers reliably track the semantics, i.e. assign to new sentences the meaning that they independently have, then the argument from novelty shares the two main weaknesses with the learnability argument.

4.3 Productivity

According to the pure argument from productivity, we need an explanation of why we are able to produce infinitely many meaningful sentences, and compositionality offers the best explanation. Classically, productivity is appealed to by Noam Chomsky as an argument for generative grammar. One of the passages runs

The most striking aspect of linguistic competence is what may call the ‘creativity of language’, that is, the speaker’s ability to produce new sentences that are immediately understood by other speakers although they

\footnote{27It is enough that the function gives meaning to sentences that have not been used, so that the speakers know the meaning of those sentence before actually using them.}
bear no physical resemblance to sentences that are ‘familiar’. The funda-
mental importance of this creative aspect of normal language use has
been recognized since the seventeenth century at least, and it was the core
of Humboldtian general linguistics (Chomsky 1971, 74).

This passage does not appeal to pure productivity, since it makes an appeal
to the understanding by other speakers (cf. Chomsky 1980, 76-78). The pure
productivity aspect has been emphasized by Fodor (e.g. Fodor 1987, 147-48),
i.e. that natural language can express an open-ended set of propositions.

However, the pure productivity argument is very weak. On the premise that
a human speaker can think indefinitely many propositions, all that is needed is
to assign those propositions to sentences. The assignment does not have to be
systematic in any way, and all the syntax that is needed for the infinity itself
is simple concatenation. Unless the assignment is to meet certain conditions,
productivity requires nothing more than the combination of infinitely many
propositions and infinitely many expressions.

4.4 Systematicity

A related argument by Fodor (Fodor 1987, 147-50) is that of systematicity. It
can be stated either as a property of speaker understanding or as an expressive
property of a language. Fodor tends to favor the former (since he is ultimately
concerned with the mental). In the simplest case, Fodor points out that if a
speaker understands a sentence of the form \textit{tRu}, she will also understand the

Formally, the argument is to be generalized to cover the understanding of
any new sentence that is formed by recombination of constituents that occur
and construction forms that are used in sentences already understood. Hence, in
this form it reduces to one of three different arguments; either to the argument
from novelty, or to the productivity argument, or finally, to the argument from
intersubjectivity (below), and only spells out a bit the already familiar idea of
old parts in new combinations.

It might be taken to add an element, for it not only aims at explaining the
understanding of new sentences that is in fact manifested, but also predicts
what new sentences will be understood. However, Fodor himself points out the
problem with this aspect, for if there is a sentence \textit{s} formed by a recombination
that we do not find meaningful, we will not take it as a limitation of the systematicity of our understanding, but as revealing that the sentence \textit{s} is not in fact
meaningful, and hence that there is nothing to understand. Hence, we cannot
come to any other conclusion than that the systematicity of our understanding
is maximal.

The systematicity argument can alternatively be understood as concerning
natural language itself, namely as the argument that sentences formed by gram-
matical recombination are meaningful. It is debatable to what extent this really
holds, and sentences (or so-called sentences) like Chomsky’s \textit{Colorless green}
ideas sleep furiously have been used to argue that not all grammatical sentences are meaningful.

But even if we were to find meaningful all sentences that we find grammatical, this does not in itself show that compositionality, or any kind of systematic semantics, is needed for explaining it. If it is only a matter of assigning some meaning or other, without any further condition, it would be enough that we can think new thoughts and have a disposition to assign them to new sentences.

4.5 Induction on synonymy

We can observe that our synonymy intuitions conform to Subst(≡µ). In case after case, we find the result of substitution synonymous with the original expression, if the new part is taken as synonymous with the old. This forms the basis of an inductive generalization that such substitutions are always meaning preserving. In contrast to the argument from novelty, where the idea of tracking the semantics is central, this induction argument may concern our habits of assigning meaning to, or reading meaning into, new sentences: we tend to do it compositionally.

There is nothing wrong with this argument, as far as it goes, beyond what is in general problematic with induction. It should only be noted that the conclusion is weak. Typically, arguments for compositionality aim at the conclusion that there is a systematic pattern to the assignment of meaning to new sentences, so that it is determined in advance what sentences so far unused mean, or will mean once they are used. This is not the case in the synonymy induction argument, for the conclusion is compatible with the possibility that the semantic function is not even defined for new sentences; the speakers only tend to extend it (cf. the learnability argument above) to new functions, just never in violation of substitutivity. It is also compatible with the possibility that the semantic function, even if it is defined for unused sentences, is not computable, and or even finitely specifiable. So, although the argument may be empirically sound, it does not establish what arguments for compositionality usually aim at.

4.6 Intersubjectivity and communication

The problems with the idea of tracking semantic when interpreting new sentences can be eliminated by bringing in intersubjective agreement in interpretation. For by our common sense standards of judging whether we understand sentences the same way or not, there is overwhelming evidence (e.g. from discussing broadcast news reports) that in an overwhelming proportion of cases, speakers of the same language interpret new sentences similarly. This convergence of interpretation, far above chance, does not presuppose that the sentences heard were meaningful before they were used. The phenomenon needs an explanation, and it is reasonable to suppose that the explanation involves the hypothesis that the meaning of the sentences are computable, and so it isn’t left to guesswork or mere intuition what the new sentences mean.
The appeal to intersubjectivity disposes of an unjustified presupposition about semantics, but two problems remain. First, when encountering new sentences, these are almost invariably produced by a speaker, and the speaker has intended to convey something by the sentence, but the speaker hasn’t interpreted the sentence, but fitted it to an antecedent thought. Secondly, we have an argument for computability, but not for compositionality.

The first observation indicates that it is at bottom the success rate of linguistic communication with new sentences that gives us a reason for believing that sentences are systematically mapped on meanings. This was the point of view in Frege’s famous passage from the opening of ‘Compound Thoughts’:

> 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 thoughts corresponding to the parts of a sentence, so that the structure of the sentence serves as the image of the structure of the thought. (Frege 1923, p. 55)

As Frege depicts it here, the speaker is first entertaining a new thought, or proposition, finds a sentence for conveying that proposition to a hearer, and by means of that sentence the hearer comes to entertain the same proposition as the speaker started out with. Frege appeals to semantic structure for explaining how this is possible. He claims that the proposition has a structure that mirrors the structure of the sentence (so that the semantic relation may be an isomorphism), and goes on to claim that without this structural correspondence, communicative success with new propositions would not be possible.

It is natural to interpret Frege as expressing a view that entails that compositionality holds as a consequence of the isomorphism idea. The reason Frege went beyond compositionality (or homomorphism, which does not require a one-one relation) seems to be an intuitive appeal to symmetry: the speaker moves from proposition to sentence, while the hearer moves from sentence to proposition. An isomorphism is a one-one relation, so that each relatum uniquely determines the other.

Because of synonymy, a sentence that expresses a proposition in a particular language is typically not uniquely determined within that language by the proposition expressed. Still, we might want the speaker to be able to work out what expression to use, rather searching around for suitable sentences by interpreting candidates one after the other. The inverse functional compositionality principle, InvFunct(µ), of section 3.4, offers such a method. Inverse compositionality is also connected with the idea of structured meanings, or thoughts,

---

28 There is some work to do to spell out the isomorphism idea, partly because one has to take account of occurrences, both of concepts in other concepts, and of expressions in other expression, partly because there is synonymy on the language side as opposed to identity on the meaning side, and partly because we need a semantic relation for the entire language, not just one separate isomorphic relation for each sentence.
while compositionality by itself isn’t, and so in this respect Frege is vindicated.\textsuperscript{29}

4.7 Summing up

Although many share the feeling that there is “more than a kernel of truth” (cf. section 4.1) in the usual arguments for compositionality, some care is required to formulate and evaluate them. One must avoid question-begging presuppositions; for example, if a presupposition is that there is an infinity of propositions, the argument for \textit{that} had better not be that standardly conceived natural or mental languages allow the generation of such an infinite set. Properly understood, the arguments can be seen as inferences to the best explanation, which is a respectable but somewhat problematic methodology. (One usually hasn’t really tried many other explanations than the proposed one.) Another important (and related) point is that standard arguments only justify the principle the meaning is computable or recursive, and the principle that up to certain syntactic variation, an expression of a proposition is computable from that proposition. Why should the semantics also be compositional, and possibly inversely compositional? One reason could be that compositional semantics, or at least certain simple forms of compositional semantics, is very \textit{simple}, in the sense that a minimal number of processing steps are needed by the hearer for arriving at a full interpretation (or, for the speaker, a full expression), but these issues of complexity have not yet been fully investigated.\textsuperscript{30}

5 Arguments against compositionality

Arguments against compositionality of natural language can be divided into four main categories:

a) arguments that certain constructions are counterexamples and makes the principle false,

b) arguments that compositionality is an empirically vacuous, or alternatively trivially correct, principle,

c) arguments that compositional semantics is not \textit{needed} to account for actual linguistic communication,

d) arguments that actual linguistic communication is not \textit{suited} for compositional semantics.

The first category, that of counterexamples, will be treated in a separate section dealing with a number of problem cases. Here we shall discuss arguments in the last three categories.

\textsuperscript{29}These ideas are developed in Pagin 2003a.

\textsuperscript{30}The conclusion of Pagin unpublished is that the hypothesis is in fact true.
5.1 Vacuity and triviality arguments

Vacuity. Some claims about the vacuity of compositionality in the literature are based on mathematical arguments. For example, Zadrozny in Zadrozny 1994 shows that for every semantics $\mu$ there is a compositional semantics $\nu$ such that $\nu(t)(t) = \mu(t)$ for every expression $t$, and uses this fact to draw a conclusion of that kind. But note that the mathematical fact is itself trivial: let $\nu(t) = \mu$ for each $t$ and the result is immediate from (2) in section 3.1 above.\footnote{Other parts of Zadrozny's results use non-wellfounded sets and are less trivial.}

Claims like these tend to have the form: for any semantics $\mu$ there is another semantics $\nu$ which is compositional and from which $\mu$ can be easily recovered. But this too is completely trivial as it stands: if we let $\nu(t) = \langle \mu(t), t \rangle$ then $\nu$ is 1-1, hence compositional by observation (3) in section 3.1, and $\mu$ is clearly recoverable from $\nu$.

In general, it is not enough that the old semantics can be computed from the new compositional semantics: for the new semantics to have any interest it must agree with the old one in some suitable sense. As far as we know there are no mathematical results showing that such a compositional alternative can always be found (see Westerståhl 1998 for further discussion).

Triviality. Paul Horwich (e.g. in Horwich 1998) has argued that compositionality is not a substantial property of a semantics, but is trivially true. He exemplifies the view with the sentence *dogs barks*, and says (Horwich 1998, 156-57) that the meaning property

(7) $x$ means DOGS BARK

consists in the so-called construction property

(8) $x$ results from putting terms whose meanings are DOG and BARK, in that order, into a schema whose meaning is NS V.

As far as it goes, the resulting semantics is compositional, and it is a trivial consequence of Horwich’s conception of meaning properties. Horwich’s view here is equivalent to Carnap’s conception of synonymy as intensional isomorphism. Neither allows that that an expression with different structure or composed from parts with different meanings could be synonymous with an expression that means DOGS BARK. However, for supporting the conclusion that compositionality is trivial, these synonymy conditions must themselves hold trivially, and that is simply not the case.

5.2 Superfluity arguments

Mental processing. Stephen Schiffer (Schiffer 1987) has argued that compositional semantics, and public language semantics altogether, is superfluous in the account of linguistic communication. All that is needed is to account for how the hearer maps his mental representation of an uttered sentence on
Schiffer argues that this can be achieved by means of transformations between sentences in Harvey’s neural language $M$. $M$ contains a counterpart $\alpha$ to (9), such that $\alpha$ gets tokened in Harvey’s so-called belief box when he has the belief expressed by (9). By an inner mechanism the tokening of $\alpha$ leads to the tokening of $\beta$, which is Harvey’s $M$ counterpart to (10). For this to be possible for any sentence of the language in question, Harvey needs a translation mechanism that implements a recursive translation function $f$ from sentence representations to meaning representations. Once such a mechanism is in place, we have all we need for the account, according to Schiffer.

The problem with the argument is that the translation function $f$ by itself tells us nothing about communicative success. By itself it just correlates neural sentences of which we know nothing except for their internal correlation. We need another recursive function $g$ that maps the uttered sentence Some snow is white on $\alpha$, and a third recursive function $h$ that maps $\beta$ on the proposition that some snow is white, in order to have a complete account. But then the composed function $h(f(g(\ldots)))$ seems to be a recursive function that maps sentences on meanings.\footnote{For an extended discussion, see Pagin 2003b.}

**Pragmatic composition.** According to François Recanati (Recanati 2004), word meanings are put together in a process of *pragmatic composition*. That is, the hearer takes word meanings, syntax and contextual features as the his input, and forms the interpretation that best corresponds to them. As consequence, semantic compositionality is not needed for interpretation to take place.

A main motivation for Recanati’s view is the ubiquity of those pragmatic operations that Recanati calls *modulations*,\footnote{These, under varying terms and conceptions, have been described e.g. by Dan Sperber and Deirdre Wilson (Sperber and Wilson 1992), Kent Bach (Bach 1994), Robyn Carston (Carston 2002) and by Recanati himself.} and which intuitively contribute to “what is said”, i.e. to communicated content before any conversational implicatures. To take an example from Recanati, in reply to an offer of something to eat, the speaker says

\begin{equation}
(11) \text{ I have had breakfast}
\end{equation}

thereby saying that she has had breakfast in the morning of the *day of utterance*, which involves a modulation of the more specific kind Recanati calls *free enrichment*, and implicating by means of what she says that she is not hungry. On Recanati’s view, communicated contents are always or virtually always prag-
matically modulated. Moreover, modulations in general do not operate on a complete semantically derived proposition, but on conceptual constituents.34 Hence, it seems that what the semantics delivers does not feed into the pragmatics.

However, if meanings, i.e. the outputs of the semantic function, are structured entities, in the sense specified by (Rev) and InvFunct(µ) of section 3.4, then the last objection is met, for then semantics is able to deliver the arguments to the pragmatic operations, e.g. properties associated with VP:s. Moreover, the modulations that are in fact made appear to be controlled by a given semantic structure: as in (11), the modulated part is of the same category and occupies the same slot in the overall structure as the semantically given argument that it replaces. This provides a reason for thinking that modulations operate on a given (syntactically induced) semantic structure, rather than on pragmatically composed material.35

5.3 Unsuitability arguments

According to a view that has come to be called radical contextualism, truth evaluable content is radically underdetermined by semantics, i.e. by literal meaning. That is, no matter how much a sentence is elaborated, something needs to be added to its semantic content in order to get a proposition that can be evaluated as true or false. Since there will always be indefinitely many different ways of adding, the proposition expressed by means of the sentence will vary from context to context.36 A characteristic example from Charles Travis (Travis 1985, 197) is the sentence

(12) Smith weighs 80 kg

Although it sounds determinate enough at first blush, Travis points out that it can be taken as true or as false in various contexts, depending on what counts as important in those contexts. For example, it can be further interpreted as being true in case Smith weighs

(12a) 80 kg when stripped in the morning
(12b) 80 kg when dressed normally after lunch
(12c) 80 kg after being force fed 4 liters of water
(12d) 80 kg four hours after having ingested powerful diuretic
(12e) 80 kg after lunch adorned in heavy outer clothing

Although the importance of such examples is not to be denied, their significance for semantics is less clear. It is in the spirit of radical contextualism to minimize the contribution of semantics (literal meaning) for determining

34In (11) it is the property of having breakfast that is modulated into having breakfast this day, not the proposition as a whole or even the property of having had breakfast.
35This line of reasoning is elaborated in Pagin and Pelletier 2007.
36Well-known proponents of radical contextualism include John Searle (e.g. Searle 1978), Charles Travis (e.g. Travis 1985), and Sperber and Wilson Sperber and Wilson 1992.
expressed content, and thereby the importance of compositionality. However, strictly speaking, the truth or falsity of the compositionality principle for natural language is orthogonal to the truth or falsity of radical contextualism. For whether the meaning of a sentence \( s \) is a proposition or not is irrelevant to the question whether that meaning is determined by the meaning of the constituents of \( s \) and their mode of composition. The meaning of \( s \) may be unimportant but still compositionally determined.

In an even more extreme version, the (semantic) meaning of sentence \( s \) in a context \( c \) is what the speaker uses \( s \) to express in \( c \). In that case meaning itself varies from context to context, and there is no such thing as an invariant literal meaning. Not even the extreme version need be in conflict with compositionality (extended to context dependence), since the substitution properties may hold within each context by itself. Context shift failure, in the sense of section 3.7, may occur, if e.g. word meanings are invariant but the meanings of complex expressions vary between contexts.

It is a further question whether radical contextualism itself, in either version, is a plausible view. It appears that the examples of contextualism can be handled by other methods, e.g. by appeal to pragmatic modulations mentioned in section 5.2 (cf. Pagin and Pelletier 2007), which does allow propositions to be semantically expressed. Hence, the case for radical contextualism is not as strong as it may \textit{prima facie} appear. On top, radical contextualism tends to make a mystery out of communicative success.

6 Problem cases

A number of natural language constructions present apparent problems for compositional semantics. In this concluding section we shall briefly discuss a few of them, and mention some others.

6.1 Belief sentences

Belief sentences offer difficulties for compositional semantics, both real and merely apparent. At first blush, the case for a counterexample against compositionality seems very strong. For in the pair

\[(13) \quad \begin{array}{l}
a. \text{John believes that Fred is a child doctor.} \\
b. \text{John believes that Fred is a pediatrician.}
\end{array}
\]

(13a) may be true and (13b) false, despite the fact that \textit{child doctor} and \textit{pediatrician} are synonymous. If truth value is taken to depend only on meaning and on extra-semantic facts, and the extra-semantic facts as well as the meanings of the parts and the modes of composition are the same between the sentences, then the meaning of the sentences must nonetheless be different, and hence compositionality fails. This conclusion has been drawn by Jeff Pelletier (Pelletier 1994).
What would be the reason for this difference in truth value? When cases such as these come up, the reason is usually that there is some kind of discrepancy in the understanding of the attributee (John) between synonyms. John may e.g. erroneously believe that pediatrician only denotes a special kind of child doctors, and so would be disposed to assent to (13a) but dissent from (13b).\textsuperscript{37} This is not a decisive reason, however, since it is what the words mean in the sentences, e.g. depending on what the speaker means, that is relevant, not what the attributee means by those words. The speaker contributes with words and their meanings, and the attributee contributes with his belief contents. If John’s belief content matches the meaning of the embedded sentence Fred is a pediatrician, then (13b) is true as well, and the problem for compositionality is disposed of.

A problem still arises, however, if belief contents are more fine-grained than sentence meanings, and words in belief contexts are somehow tied to these finer differences in grain. For instance, as a number of authors have suggested, perhaps belief contents are propositions under modes of presentation.\textsuperscript{38} It may then be that different but synonymous expressions are associated with different modes of presentation. In our example, John may believe a certain proposition under a mode of presentation associated with child doctor but not under any mode of presentation associated with pediatrician, and that accounts for the change in truth value.

In that case, however, there is good reason to say that the underlying form of a belief sentence such as (13a) is something like

\begin{equation}
\text{Bel}(\text{John, the proposition that Fred is a child doctor, M('Fred is a child doctor')})
\end{equation}

where M(-) is a function from a sentence to a mode of presentation or a set of modes of presentation. In this form, the sentence Fred is a pediatrician occurs both used and mentioned (quoted), and in its used occurrence, child doctor may be replaced by pediatrician without change of truth value. Failure of substitutivity is explained the fact the surface form fuses a used and a mentioned occurrence. In the underlying form, there is no problem for compositionality, unless caused by quotation.

Of course, this analysis is not obviously the right one, but it is enough to show that the claim that compositionality fails for belief sentences is not so easy to establish.

6.2 Quotation

Often quotation is set aside for special treatment as an exception to ordinary semantics, which is supposed to concern used occurrences of expressions rather than mentioned ones. Sometimes, this is regarded as cheating, and quotation is proposed as a clear counterexample to compositionality: brother and male

\textsuperscript{37}Cf. Benson Mates (Mates 1950) and Tyler Burge Burge 1978. Mates took such cases as a reason to be skeptical about synonymy.

\textsuperscript{38}See e.g. Burdick 1982, Salmon 1986. Salmon, however, existentially quantifies over modes of presentations, which preserves substitutivity.
sibling are synonymous, but ‘brother’ and ‘male sibling’ are not (i.e. the expressions that include the opening and closing quote). Since enclosing an expression in quotes is a syntactic operation, we have a counterexample.

Let us assume quoting is a genuine syntactic operation. Then the syntactic rules include a total unary operator \( \kappa \) that applies to expressions to form expressions enclosed in quotes. That is, for any simple or complex expression \( e \),

\[
\kappa(e) = 'e'
\]

(i.e. the string leftquote \( e \) rightquote). The semantics of quoted expressions is given simply by

\[
(Q) \quad \mu(\kappa(t)) = V(t)
\]

for any meaningful term \( t \). (Recall from section 2 the string value function \( V \), and the convention to use \( \pi \) as a name for \( \kappa \).) \( Q \) gives the semantic value of quote-terms: it is simply the string value of the argument to the quote operator. Then, since \( t \equiv_\mu u \) does not imply \( V(t) = V(u) \), substitution of \( u \) for \( t \) in \( \pi(t) \) may violate compositionality.

However, not much is required for transforming such a non-compositional semantics for quotation into a compositional one. To see how this can be done, start with a grammar \( E = (E, A, \Sigma) \) (for a fragment of English, say) and a compositional semantics \( \mu \) for \( E \). First, extend \( E \) to a grammar containing the quotation operator \( \kappa \). To describe this in detail one would need to know more about \( E \), but the idea is simple: we need to allow not only quote-strings of the form ‘John’, ‘likes’, “Mary”, etc., but also things like John likes ‘Mary’ (meaning that he likes the word), whereas we should disallow things like John ‘likes’ Mary (this is not an account of scare quotes) or ‘John likes’ Mary as ungrammatical. This involves extending each function \( \alpha \in \Sigma \) to a function \( \alpha' \) in a way which accounts for such facts; the details should be fairly straightforward (for example, one could treat all quote-strings as NPs). Let \( E' \) be the closure of \( E \) under these operations and \( \kappa \) (that we take to be a total unary function on \( E' \)), and let \( \Sigma' = \{ \alpha' : \alpha \in \Sigma \} \cup \{ \kappa \} \). Then we have a new grammar \( E' = (E', A, \Sigma') \) that incorporates quotation.

Second, there is an almost equally straightforward way to extend \( \mu \) to a semantics \( \mu' \) for \( E' \). Since \( \mu \) satisfies Funct(\( \mu \)), there corresponds a semantic operation \( r_\alpha \) to each \( \alpha \in \Sigma \), and this is extended to an operation \( r_{\alpha'} \), roughly by including the strings of \( E' \) among the objects of the universe, with certain properties and certain relations to other objects. Also, use the semantics \( Q \) above for \( \kappa \).

As we already indicated, the straightforward semantics \( \mu' \) is not compositional: even if Mary is the same person as Sue, John likes ‘Mary’ doesn’t mean the same as John likes ‘Sue’ . However, we now show how to extend \( \mu' \) to a semantics \( \mu'' \) for \( E' \) that is compositional is the sense of LC-Funct(\( \mu \)) in section 3.7, by adapting Frege’s view in Frege 1892 that quotation provides a context
Thus, assume there are two context types: \(c_u\), the use context type, which is the default and is the type for the null context (for unembedded occurrences), and \(c_q\), which is the quotation context type. The function \(\psi\) from context types and operators to context types is given by

\[
\psi(c, \beta) = \begin{cases} 
c & \text{if } \beta \neq \kappa \\
c_q & \text{if } \beta = \kappa
\end{cases}
\]

for \(\beta \in \Sigma'\) and \(c\) equal to \(c_u\) or \(c_q\).

Let \(M' = \text{range}(\mu')\). We can assume \(M' = M \cup E'\), where \(M\) and \(E'\) are disjoint and the elements of \(E'\) only serve as meanings of quote-terms. First, for an atomic term \(a \in A\),

(i) \(\mu''(a, c) = \begin{cases} 
\mu'(a) (= \mu(a)) & \text{if } c = c_u \text{ and } \mu(a) \text{ is defined} \\
a & \text{if } c = c_q
\end{cases}
\)

(undefined otherwise). To specify the meaning of complex terms in contexts, we define new operations \(r''_{\alpha'}\) for \(\beta \in \Sigma'\). If \(\alpha \in \Sigma\) is \(n\)-ary and \(m_1, \ldots, m_n \in M'\) let

\[
r''_{\alpha'}(m_1, \ldots, m_n, c) = \begin{cases} 
 r_{\alpha'}(m_1, \ldots, m_n) & \text{if } c = c_u, \ m_1, \ldots, m_n \in M, \text{ and } r_{\alpha'} \text{ is defined for } m_1, \ldots, m_n \\
 \alpha'(m_1, \ldots, m_n) & \text{if } c = c_q, \ m_1, \ldots, m_n \in E', \text{ and } \alpha' \text{ is defined for } m_1, \ldots, m_n \\
 \text{undefined} & \text{otherwise}
\end{cases}
\]

That is, when the term occurs in a use context, \(r''_{\alpha'}\) reduces to \(r_{\alpha'}\), and in a quotation context it reduces to the syntactic operator \(\alpha'\) itself. Next, if \(\beta = \kappa\) and \(m \in M'\), let

\[
r''_{\kappa}(m, c) = \begin{cases} 
m & \text{if } c = c_u \text{ and } m \in E' \\
\kappa(m) & \text{if } c = c_q \text{ and } m \in E' \\
\text{undefined} & \text{if } m \in M
\end{cases}
\]

Now we can complete the definition of \(\mu''\). For any \(\beta \in \Sigma'\) and \(t_1, \ldots, t_n \in GT_E\), if \(c\) is a context type and \(c' = \psi(c, \beta)\), let

\[39\textit{Markus Werning Werning 2005 gives a compositional semantics by treating all apparent quotation of non-atomic strings as concatenation of names of atomic strings (letters, phonemes), and by treating those names as unstructured primitives. Thus, he does not have a quotation operator. Christopher Potts Potts 2007 gives a semantics for quotation with a quote function similar to }\kappa\textit{, although he works with expression triples (roughly) like those in section 3.6, and lets the quote function apply to such triples. But, in spite of his (at least implicit; cf. Potts 2007, pp. 406, 426) claim to the contrary, this semantics is not compositional (although it is recursive; cf. section 3.2), for precisely the reasons exemplified here: if Mary is Sue, }\textit{Mary' and 'Sue' are synonymous in Potts' semantics, but 'John likes 'Mary' and 'John likes 'Sue' are not. (More exactly, this holds for the corresponding triples.) We venture the conjecture that an adequate compositional treatment of quotation needs something like the contextual parameters introduced below.}
(ii) \( \mu''(\overline{\tau}(t_1, \ldots, t_n), c) = r''_{\gamma}(\mu''(t_1, c'), \ldots, \mu''(t_n, c'), c) \)

whenever the right-hand side is defined (undefined otherwise). We then have the following:

(15) \( \text{LC-Funct}(\mu'')_c \) holds.
(16) \( \mu'' \) extends \( \mu \) in the sense that if \( t \in GT_E \) is meaningful, \( \mu''(t, c_u) = \mu(t) \).
(17) For all meaningful \( t \in GT_E \), \( \mu''(\overline{\tau}(t), c_u) = \mu''(t, c) = V(t) \).

So \( \mu'' \) is compositional in the contextually extended sense. That \( t \equiv_{\mu} u \) holds does not license substitution of \( u \) for \( t \) in \( \overline{\tau}(t) \), since \( t \) there occurs in a quotation context, and we may have \( \mu''(t, c_q) \neq \mu''(u, c_q) \).

### 6.3 Idioms

Idioms are almost universally thought to constitute a problem for compositionality. Usually the thought here is of the following kind: The VP kick the bucket can also mean DIE, but the semantic operation corresponding to the standard syntax of, say, fetch the bucket, giving its meaning in terms of the meanings of its immediate constituents fetch and the bucket, cannot be applied to give the idiomatic meaning of kick the bucket.

This is no doubt a problem of some sort, but not necessarily for compositionality. First, that a particular semantic operation fails doesn’t mean that no operation works, but just that would be required to violate compositionality. Second, we must take into account that kick the bucket is actually ambiguous between its literal and its idiomatic meaning, but compositionality presupposes non-ambiguous meaning bearers. Unless we take the ambiguity itself to be a problem for compositionality (about this, see the next subsection), we should first find a suitable way to disambiguate the phrase, and only then raise the issue of compositionality.

Such disambiguation may be achieved in various ways. We could treat the whole phrase as a lexical item (an atom), in view of the fact that its meaning has

\[ \mu''(\overline{\tau}(t_1, \ldots, t_n), c) = r''_{\gamma}(\mu''(t_1, c'), \ldots, \mu''(t_n, c'), c) \]

\[ \mu''(t, c_u) = \mu(t) \]

\[ \mu''(\overline{\tau}(t), c_u) = V(t) \]

\[ t \equiv_{\mu} u \]

\[ \mu''(t, c_q) \neq \mu''(u, c_q) \]

To see how the semantic rules work, suppose the string John likes Mary is the value of the term \( t_1 = \overline{\tau}(\text{John}, \overline{\tau}(\text{like}, \overline{\tau}(\text{Mary})) \) in the given grammar, and that its meaning \( \mu(t_1) \) is \( r_\gamma(j, r_{\gamma'}(\text{LIKE}, m)) \). Then John likes 'Mary' will be the value of the term \( t_2 = \overline{\tau}(\text{John}, \overline{\tau}(\text{like}, \overline{\tau}(\text{Mary}))) \) in \( E' \), whose meaning is calculated as follows:

\[ \mu''(t_2, c_u) = r''_{\gamma}(\mu''(\text{John}, c_u), \mu''(\overline{\tau}(\text{like}, \overline{\tau}(\text{Mary})), c_u)) \]

\[ = r''_{\gamma}(j, r_{\gamma'}(\mu''(\text{like}, c_u), \mu''(\overline{\tau}(\text{Mary}, c_u), c_u))) \]

\[ = r''_{\gamma}(j, r_{\gamma'}(\text{LIKE}, r''_{\gamma}(\mu''(\text{Mary}, c_u), c_u))) \]

\[ = r''_{\gamma}(j, r_{\gamma'}(\text{LIKE}, r''_{\gamma}(\text{Mary}, c_u), c_u)) \]

\[ = r''_{\gamma}(j, r_{\gamma'}(\text{LIKE}, \text{Mary}, c_u)) \]

\[ = r''_{\gamma}(j, r_{\gamma'}(\text{LIKE}, \text{Mary})) \]

\[ i.e., \text{just like the meaning of } t_1, \text{ except that } \text{Mary} \text{ (the woman) is replaced by } \text{Mary} \text{ (the word). Similarly, one calculates that } \mu''(t_2, c_q) \text{ is the string } \text{John likes 'Mary'}. \]
to be learnt separately. Or, given that it does seem to have syntactic structure, we could treat it as formed by a different rule than the usual one. In neither case is it clear that compositionality would have a problem.

To see what idioms really have to do with compositionality, it is convenient to think of the following situation. Given a grammar and a compositional semantics for it, suppose we decide to give some already meaningful phrase a non-standard, idiomatic meaning. Can we extend the given syntax (in particular, to disambiguate) and semantics in a natural way that preserves compositionality? Note that it is not just a matter of accounting for one particular phrase, but rather for all the phrases in which the idiom may occur. This requires an account of how the syntactic rules apply to the idiom, and to its parts if it has structure, as well as a corresponding semantic account.

Interestingly, idioms differ as to syntactic behavior. While the idiomatic kick the bucket is fine in *John kicked the bucket yesterday*, or *Everyone kicks the bucket at some point*, it is not fine in

(18) The bucket was kicked by John yesterday.
(19) Andrew kicked the bucket a week ago, and two days later, Jane kicked it too.

In contrast, pull strings preserves its idiomatic meaning in passive form, and strings is available for anaphoric reference with the same meaning:

(20) Strings were pulled to secure Henry his position.
(21) Kim’s family pulled some strings on her behalf, but they weren’t enough to get her the job.

This suggests that these two idioms should be analyzed differently; indeed the latter kind is called “compositional” in Nunberg, Sag, and Wasow 1994 (from which (21) is taken), and is analyzed there using the ordinary syntactic and semantic rules for phrases of this form but introducing instead idiomatic meanings of its parts (*pull* and *string* in this case), whereas kick the bucket is called “non-compositional”.

In principle, however, nothing prevents a semantics taking care of the two kinds of idioms in different ways from being compositional in our sense. Incorporating idioms in syntax and semantics is an interesting task. For example, in addition to explaining the facts noted above one has to prevent kick the pail from meaning ‘die’ even if bucket and pail are synonymous (if one strives for compositionality, that is), and likewise to prevent the idiomatic versions of pull and string to combine illegitimately with other phrases. For an overview of the semantics of idioms, see Nunberg, Sag, and Wasow 1994. Westerståhl 2002 is an abstract discussion of various ways to incorporate idioms while preserving

\footnote{That rule would have the same effect as the ordinary rule, concatenating the transitive verb with the NP, but it would have a different name, so another semantic operation could correspond to it. If you think this is cheating, observe that if it were the same rule and the same components, we are back to the problem of ambiguity.}
compositionality.

6.4 Ambiguity

Even though the usual formulation of compositionality requires non-ambiguous meaning bearers, the occurrence of ambiguity in language is usually not seen as a problem for compositionality. This is because lexical ambiguity seems easily dealt with by introducing different lexical items for different meanings of the same word (say, by indexing), whereas structural ambiguity corresponds to different derivation histories of the same surface string, and so if meaning is assigned to these histories (terms in the term algebra), as we have done here, the ambiguity disappears.

It is possible, however, to be dissatisfied with this analysis. More precisely, one may argue that even though there are clear cases of structural ambiguity in language, as in Old men and women were released first from the occupied building, in other cases the additional structure is just an ad hoc way to avoid ambiguity. In particular, scope ambiguities could be taken to be of this kind. For example, while semanticists since Montague have had no trouble inventing different underlying structures to account for the two readings of

\[(22) \text{ Every critic reviewed four films.}\]

it may be argued that this sentence in fact has just one structural analysis, a simple constituent structure tree, and that meaning should be assigned to that one structure. A consequence is that meaning assignment is no longer functional, but relational, and hence compositionality either fails or is just not applicable. Pelletier 1999 draws precisely this conclusion.

Even if one agrees with such an account of the syntax of (22), abandonment of compositionality is not the only option, however. One possibility is to give up the idea that the meaning of (22) is a proposition, i.e. something with a truth value (in the actual world), and opt instead for underspecified meanings of some kind. Such meanings can be uniquely, and perhaps compositionally, assigned to simple structures like constituent structure trees, and one can suppose that some further process of interpretation of particular utterances leads to one of the possible specifications, depending on various circumstantial facts. This is a form of context-dependence, and we saw in section 3.7 how similar phenomena can be dealt with compositionally. What was there called standing meaning is one kind of underspecified meaning, represented as a function from indices to ‘ordinary’ meanings. In the present case, where just a few meanings, all specifiable in advance, are available, one might try to use the set of those meanings instead. A similar but more sophisticated way of dealing with quantifier scope is so-called Cooper storage (see Cooper 1983). It should be noted, however, that while such strategies restore a functional meaning assignment, the compositionality of the resulting semantics is by no means automatic; it is an issue that has to be addressed anew.

If one dislikes the idea that ambiguity leads to underspecified meanings, an-
other option might be to accept that meaning assignment becomes relational and attempt instead to reformulate compositionality for such semantics. Although this line has hardly been tried in the literature, it may be an option worth exploring.  

6.5 Other problems


All in all, it seems that the issue of compositionality in natural language will remain live, important and controversial for a long time to come.

References


42For some first attempts in this direction, see Westerståhl forthcoming Yet another possible route to deal with ambiguity while preserving compositionality was hinted at in section 3.6; cf. footnote 24.


